Contextuality and Ontological Models:
A Tale of Desire and Disappointment

by

Piers Lillystone

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Physics (Quantum Information)

Waterloo, Ontario, Canada, 2019

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner: Jan-Åke Larsson
                        Professor, Head of the Dept. of Electrical Engineering,
                        Dept. of Electrical Engineering, Linköping University

Supervisor: Joseph Emerson
                        Professor, Dept. of Applied Mathematics, University of Waterloo

Internal Member: Kevin Resch
                        Professor, Dept. of Physics and Astronomy, University of Waterloo

Internal-External Member: Robert Spekkens
                        Adjunct Assistant Professor, Dept. of Physics and Astronomy,
                        University of Waterloo

Other Member(s):   Richard Cleve
                        Professor, Dept. of Computer Science, University of Waterloo

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Since being defined by Kochen and Specker, and separately by Bell, contextuality has been proposed as one of the key phenomena that distinguishing quantum theory from classical theories. However, with the rise of quantum information contextuality's position as the leading definition of the quantum/classical boundary has been called into question. This is due to the fact that a contextual explanation is required by subtheories that offer no exponential quantum computational advantage over classical computation. In this thesis, submitted in requirement for a PhD in physics with quantum information, we shall explore this unwanted contextuality, and show that generalized contextuality is more prevalent than was previously thought.

First we will show that the single-qubit stabilizer subtheory, which was previously thought of as a non-contextual subtheory, requires a generalized contextual ontological model, when transformations are included in the operational description. In addition to this we show that the even smaller single-rebit stabilizer subtheory, a strict subtheory of the single-qubit stabilizer subtheory, must also admit a generalized contextual ontological model, again when transformations are included in the operational description. We then show that both these subtheories require negatively represented quasi-probability representations, re-establishing the link between contextuality and negativity for transformations. We also investigate the representation of transformations in generic quasi-probability representations, showing that under a reasonable assumption almost-all unitaries must be negatively represented by a finite quasi-probability representation.

Second we will investigate the efficiently simulable n-qubit stabilizer subtheory, which exhibits all forms of contextuality and thus represents the main obstacle to identifying contextuality as a resource for quantum computation. To this end we present an attempt at constructing a model based on a frame-switching Wigner function. This leads us to constructing a contextual ψ-epistemic ontological model of the n-qubit stabilizer formalism. We shall see that this model is outcome deterministic, which is one of the core assumptions in the definition of traditional non-contextuality. This model then will lead us to a result that proves that any ontological model of the n-qubit stabilizer formalism requires at least $n-1$ generators to be encoded in the ontology of the model. As $n-1$ generators represents almost full knowledge about the stabilizer state, we therefore posit that ψ-onicity, a more controversial notion of non-classicality, is actually *the* resource for quantum computation.

Acknowledgements

First and foremost this thesis would not have been possible without the support from my supervisor, Professor Joseph Emerson. Thank you for having faith in me, and allowing me the opportunity to realize my dream to work at the forefront of theoretical physics.

I'd also like to extend my thanks to Joel J. Wallman who's guidance allowed me to come to a far deeper understanding of the material presented in this thesis. Joel was also the source of many of the projects I worked on for this thesis. I will always be humbly grateful for his support and insightful comments on the projects.

The role of my PhD committee in helping my achieve this PhD can also not be under-stated, and I'd like to thank Richard Cleve, Joseph Emerson, Kevin Resch, and Robert Spekkens for the all great feedback and direction they provided throughout the years. In particular I'd like to thank Robert Spekkens for his expertise on the subjects covered in this thesis.

My deepest gratitude is also extended to all the members of our research group, who were a constant pleasure to work with. Specifically I would like to thank Sumit Sijher for being the best companion I could ask for during our adventures in academia. Also I'd like to thank Hammam Qassim, and Joshua Ruebeck for their academic insights, general brilliance and humour, and indulging my predisposition for rambling on about the foundations of physics. I'd also like to thank Arnaud Carigan-Dugas, Dmitry Serbin, Anirudh Krishna and Thom Bohdanowicz for helping me feel welcomed when I first arrived in Canada, and continue to be great friends to this day.

Within the wider academic community there are many people I'd like to thank for the stimulating conversations about research and being genuinely wonderful people to have the pleasure of meeting. Some of these people are, but not limited to, Juani Bermjo-Vega, Angela Karanjai, Lorenzo Catani, David Schmidt, Matthew Pusey, Matthew Leifer, Shane Mansfield, and Robert Raussendorf.

I also owe the institute for quantum computing a huge thanks. It has been a fantastic institute to work and study at. It has provided many opportunities that have helped not only my academic career, but my life general, and it shall always hold a dear place in my heart.

Finally I must thank all the people that supported me through this journey. Without your support none of this thesis would have been possible. Throughout it all my family have been the rock that I was able to stand on, and I owe them a debt I can never repay. Iara, your love gave me the strength to acheive my dreams, and I am excited to see where

our love will take us next. I also have to thank all my friends. You were there for me during the best and worst of times, I am continually honoured that you consider me a part of your life.

Dedication

This thesis is dedicated to Mikka and the house she loved, 9 John street. They will both be sorely missed, may they both enjoy their rest.

Table of Contents

List of Figures

xiii

PART I

Introduction

Since the advent of quantum theory, almost a century ago, there has been a long standing question concerning what quantum theory implies about the physical reality we live in. Many of physic's heavy-weights have tried to resolve this problem, from Einstein [26] to Bell [8, 7] and Kochen-Specker [43] to more modern approaches [55, 63, 36, 66, 39]. However, to this day none of these insights can fully explain the boundary between classical theories and quantum theory.

The question of what distinguishes quantum theory from a classical theory has seen a resurgence in importance, with the rise of the field of quantum computation. A quantum computer offers the potential for vast improvements, probably exponential, over the capabilities of a classical computer. Therefore, it is critical that we understand the source of this advantage, such that more efficient and more powerful quantum computers and quantum algorithms can be constructed.

Before continuing, I'd like to take a moment to comment that this thesis, and indeed my research interests, are not concerned with taking or constructing an interpretational viewpoint on quantum theory. As it would appear to me that any attempt at providing an all-encompassing interpretation of quantum theory is doomed to failure, until we fully understand the phenomena predicted by quantum theory. That isn't to say that this thesis doesn't contain content that could be seen as constituting some kind of interpretation of quantum theory. Rather, the interpretational elements in this thesis, namely the ontological models formalism, have been adopted for their mathematical rigor and not their claims about the structure of the physical universe[1].

There have been many suggestions into what classifies the quantum/classical boundary, with non-locality [8, 7] and contextuality [43, 76] being the most widely accepted. Indeed non-locality is a special case of contextuality, where contextuality is supplemented with an assumption of space-like separation. One of the emerging viewpoints on a robust definition of the quantum/classical boundary is to define it by whether a given set of quantum processes admits an efficient classical simulation. If not the quantum processes must exhibit some phenomena that is truly quantum. Thus this phenomena can be seen as a *resource* for quantum computation. However both definitions of the quantum/classical divide provided by non-locality and contextuality fail to uniquely identify the source of a quantum computer's power. Where non-locality is a sufficient resource for quantum computation, but not necessary [16], and contextuality is a necessary resource for quantum computation, but not sufficient [39, 9].

This poses the question: If contextuality and non-locality *do not* uniquely identify

---

[1]Although having said this, in my opinion, the existence of a real physical state of a system would appear to be an appealing feature of a physical theory.

the quantum/classical boundary, what might? Investigating this question, and hopefully providing some insights into it, is the aim of this thesis. To do so we use the operational theories and ontological model framework, which provides rigorous mathematical tools that can classify entire classes of physical theories[2]. Further, this thesis is focused on subtheories of quantum theory that do not admit a non-contextual ontological model, but do not offer a quantum computational advantage over polynomial classical computation. We have restricted our investigations to these setting such that the nature of the contextuality present within these subtheories can be studied and identified. Hopefully leading to a possible refinement of the notion of contextuality that is both necessary *and* sufficient for universal quantum computation. However, one of the more surprising results presented in this thesis is that a more contentious definition of non-classicality, the requirement of a ψ-ontic ontological model of an operational theory, could be the defining line between classical theories and quantum theory.

Structure of the Thesis

This thesis has been structured in such a way that each chapter forms an independent body of text, with all relevant information to the content of the chapter presented in the chapter. This is not to say that there is no relation between each of the chapters, quite opposite, but that each chapter can be read without knowledge of content elsewhere in the thesis. I made this decision as many of the chapters are reproductions of research papers, with some additional material, and therefore are by construction self-contained. This was then extended to all additional chapters. My hope is that if an aspect of this thesis interests the reader, then they may focus solely on the chapters of interest, without reference the rest of the thesis.

Part II - Contextuality, Negativity, and Subtheories of a Single-Qubit: A Case of Mistaken Identity

In this part, I will present research into transformation contextuality in ontological models, and transformation negativity in quasi-probability representations. Both topics represent under-explored aspects of the contextuality and negativity landscape, and offer the potential for novel results. They are also more directly applicable to the circuit model of

---

[2]Although arguably not all possible theories.

quantum computation, and I believe could offer interesting insights into the resources needed for universal quantum computation.

Chapters 1 and 2 focus on the results that the single-qubit stabilizer formalism and single-rebit stabilizer formalism require a generalized contextual ontological model and a negative quasi-probability representation, when transformation are included in the operational theory. In chapter 3 I extend the scope of the previous work and present a result that in *any* finite quasi-probability representation *almost-all* unitaries are negatively represented, given a natural assumption about the representation of the identity transformation.

Chapter 1 - Contextuality and the Single-Qubit Stabilizer Formalism

In chapter 1, we shall show that the single-qubit stabilizer subtheory requires a generalized contextual ontological model, when transformations are included in the operational theory. This is contrary to the previously held view that the subtheory was a universally non-contextual subtheory of quantum theory.

The single-qubit stabilizer subtheory is possibly one of the simplest, and arguably classical as it offers no super-polynomial advantage over classical computation, subtheories of quantum theory we can construct, that is not a trivial subtheory. It is also the basic building block of many applications inside the field of quantum computation. Therefore the fact that it requires a generalized contextual ontological model is a very surprising result, considering that generalized contextuality has been considered to be one of the defining lines between classical theories and quantum theory. Hence, this result casts further doubt on whether generalized contextuality can be considered a quantum phenomena, with the n-qubit stabilizer subtheory[3] already exhibiting all forms of contextuality.

This results also demonstrates a fatal flaw in only considering prepare-measure scenarios. Contextuality can effectively be "hidden" in the transformations of the model, this can be seen by considering our result is bi-directional:

- Preparation non-contextuality $\Rightarrow$ transformation contextuality. Therefore, transformation non-contextuality $\Rightarrow$ preparation contextuality.

- Outcome determinism [4] $\Rightarrow$ transformation contextuality. Transformation non-contextuality $\Rightarrow$ outcome indeterminism.

---

[3] $n > 1$.

[4] Where we note that the assumption of outcome determinism *and* measurement non-contextuality is equivalent to Kochen-Specker (traditional) contextuality.

Thus the prepare-measure operational setting does not fully capture the contextuality of quantum theory. As an aside, I have been unable to construct a transformation non-contextual ontological model of the single-qubit stabilizer subtheory, therefore it may be possible that transformation contextuality is required regardless of the representations of states and measurements.

I also present as supplementary material some of the further research that was done after the project was completed, investigating some of the open questions the project posed. Firstly, I present an alternative proof of transformation contextuality of the single-qubit stabilizer subtheory. Unlike the proof presented in the paper, this proof uses the depolarizing channel and coarse-grained measurement to prove transformation contextuality. This proof has some desirable features that are not shared with the results in the paper. Secondly, we will investigate the link between transformation contextuality for a single-qubit and traditional contextuality for 2-qubits, via the Choi-Jamiołkowski isomorphism. We shall see that there is a qualitative link between the two, however there are still unresolved questions about how to make this connection quantitative.

Chapter 2 - Contextuality and Negativity for Subtheories of a Single Qubit

In chapter 2 we shall extend the results presented in chapter 1. We will prove that the single-rebit subtheory requires a generalized contextual representation when transformations are included in the subtheory. Additionally, we show that both the single-qubit stabilizer subtheory and single-rebit subtheory require negatively represented quasi-probability representations.

This chapter is intended to be a more technical extension of the results presented in chapter 1, and constitutes its own paper. However, this does not mean that the results contained within are any less surprising. Indeed, it was already surprising that the single-qubit stabilizer subtheory required a generalized contextual model. The single-rebit subtheories[5] are strict subtheories of the single-qubit stabilizer subtheory. So these results show that generalized contextuality is *even more* ubiquitous than the results in chapter 1 would suggest.

The more technical aspects of the chapter lie in the proofs that both the single-rebit subtheory and the single-qubit stabilizer subtheory require a negatively represented quasi-probability representation. This result begins to connect transformation contextuality to transformation negativity. Much like the previously established connection between

---

[5]There are 3 different subtheories the term "rebit" could refer too. With each one defined by the set of operators which have real representations in the X, Y, and Z bases.

contextuality and negativity in the prepare-measure setting [78, 28]. However, unlike previous results this connection is explicitly established for a subtheory of quantum theory.

Chapter 3 - Quasi-Probability Representations of Transformations

In chapter 3 we shall show that, given a plausible assumption about the representation of the identity transformation, in *any* finite quasi-probability representation *almost-all* unitaries are negatively represented.

This result is in stark contrast to the case of preparations and measurements, where we can construct a quasi-probability representation that represents either all preparations or all measurements positively. However, in its current state the result relies on the assumption that the identity transformation, or "do nothing" channel, is represented by the identity matrix. Hence, in the chapter we also explore ways to remove this assumption and prove this statement directly from the definitions of a quasi-probability representation.

If this result can be conclusively proven, without the assumption, it would reveal a very interesting property of quantum theory. Namely, that in the context of representing quantum theory, transformations are distinguished from preparations and measurements.

Part III - The n-qubit Stabilizer Subtheory and Ontological Models Thereof

In Part III, we will construct an ontological model of the n-qubit stabilizer subtheory, and then use insights from this model to make general statements about any ontological model of the stabilizer subtheory. The n-qubit stabilizer subtheory forms the cornerstone of many quantum information processing tasks, being the basis for most error-correcting techniques, and cannot confer a quantum computational advantage as there exists efficient classical simulations of the subtheory. However, from a foundational perspective the n-qubit stabilizer subtheory requires a contextual ontological model, which is usually seen as a signature of quantumness.

To resolve this mismatch in the definitions of what signifies the classical/quantum divide, I constructed a ψ-epistemic model of the n-qubit stabilizer subtheory. In chapter 4, we shall investigate the original attempt at constructing a classical model of the subtheory, which is done via a frame-switching Wigner function. This chapter also serves as an in-depth background to the theoretical machinery used in the following chapters, with

additional discussion on how our constructed ψ-epistemic model relates to the recent quasi-probability simulation by Raussendorf et al [69]. In chapter 5 we present the contextual ψ-epistemic model of the n-qubit stabilizer formalism, and investigate its properties. In chapter 6 we apply insights from constructing the model to prove that in *any* ontological model of the n-qubit stabilizer formalism at least $n-1$ generators must be encoded in the ontology. I.e. almost all the information required to specify a stabilizer state must be encoded in the ontology. This suggests a rather surprising conclusion; is ψ-onticity the actual "resource" of quantum computation?

Chapter 4 - Using the Wigner Function to Represent the n-Qubit Stabilizer Subtheory

In chapter 4, we will demonstrate the failure of a frame-switching Wigner function to reproduce the n-qubit stabilizer subtheory's statistics, which was hoped to provide an analogue to the Wigner function for qudits [82, 83, 39]. The failure of this model occurs when we try and represent state update post-measurement positively. This result inevitably leads to the results in chapters 5 and 6, and the realization that to correctly model state update post-measurement we are required to track $n-1$ generators of a stabilizer group. This was the fatal flaw in the frame-switching Wigner function, as each point in its phase-space could encode at most 1 generator.

The aims of this chapter are three-fold: To provide an in-depth introduction to the stabilizer formalism and quasi-probability representation framework, hence this chapter is more to be seen as an introductory chapter to chapters 5 and 6. To present the frame-switching Wigner function and explain how this led to the results presented in chapters 5 and 6. And finally, and most importantly, to provide the groundwork in connecting the quasi-probability simulation of Raussendorf et al to the contextual ψ-epistemic model presented in this paper.

The results presented in this chapter can be summarized by four key theorems. The first theorem, theorem 4.4.3, provides an if and only if condition on whether a stabilizer state is positively represented within a given frame. The second theorem, theorem 4.4.5, shows that we can achieve a positive representation of Clifford transformations if we couple them with a frame switching operation, even if the representation of each individual operation may be negative. The third theorem, theorem 4.4.6, proves that in any frame all stabilizer measurements have positive response functions. The fourth and most important theorem, theorem 4.4.8, proves that in any given frame there exists a Pauli measurement whose corresponding update map after measurement must be negatively represented, scuttling

any hope of using the frame-switching Wigner function to represent the n-qubit stabilizer subtheory positively.

Chapter 5 - A Contextual ψ-Epistemic Model of the n-Qubit Stabilizer Formalism

In chapter 5, I will present my construction of a contextual ψ-epistemic model of the n-qubit stabilizer subtheory. This construction, as presented in the paper, has been chosen such that outcome determinism of measurement outcomes is maintained. This choice has been made as outcome determinism is often seen as a core component of traditional contextuality. Additionally I provide a construction of the model where outcome determinism has been dropped, which significantly reduces the size of the ontology required by the model and causes the model to become a *symmetric-always-ψ*-epistemic model, which will be defined in chapter 5.

In the paper, I split presenting the model into two cases, the 2-qubit model and the n-qubit model. This has been done to slowly introduce the concepts used in the model by presenting the most simple case first, with the single-qubit model being equivalent to the 8-state model of a single qubit. Furthermore, many of the proofs of contextuality that can be constructed in the n-qubit stabilizer formalism only require 2-qubits, the most well-known being the Mermin-Peres square. Therefore, I discuss how the model reproduces the statistics of the Mermin-Peres square and comment on the assumption of contextuality that underpins the squares construction. Additionally the Pusey-Barrett-Rudolph (PBR) theorem can also be reproduced within the 2-qubit stabilizer formalism, via a sequence of two measurements where the choice of second measurement is controlled by the outcome of the first measurement. Therefore, I demonstrate how the model accounts for the PBR theorem.

To compliment the result, a finite-state machine representation of the 2-qubit model was constructed. The finite-state machine gives a classically appealing representation of the model and allows for us to graphically simulate the model. The code that performs this simulation has been included in the appendix, in a reduced form.

Chapter 6 - Is the Stabilizer Subtheory the Limit of ψ-Epistemic Models?

In chapter 6, I present a result that proves that any ontological model of the n-qubit stabilizer subtheory *must* encode $n-1$ generators of a stabilizer state's stabilizer group in the ontology of the ontological model. This result suggests that the stabilizer subtheory

is on the verge of ψ-onticity, as $n-1$ generators represent almost complete knowledge about the stabilizer state. Therefore, it is hard to see how the inclusion of any additional elements to the subtheory would not require more information to be encoded within the ontology.

This result is derived without any additional assumptions about the structure of the ontological model, and only relies on the group structure of the stabilizer subtheory. To prove this result I first prove that in any ontological model of the 2-qubit stabilizer subtheory at least one generator is required to be encoded in the ontology, then via an embedding procedure show that this result implies $n-1$ generators must be encoded in any ontological model of the n-qubit stabilizer subtheory. The proof uses a novel embedding proposition that allows us to directly work with the stabilizer groups rather than the ontological elements, reducing much of the complexity of the proof.

PART II

Contextual, Negativity, and
Subtheories of a Single-Qubit: A
Case of Mistaken Identity

Chapter 1

Contextuality and the Single-Qubit Stabilizer Formalism

1.1 Chapter Preamble

In this chapter I present work completed with Joel J. Wallman and Joseph Emerson. In the work, published in Physical Review Letters (12th April 2019) [52], we showed that the single-qubit stabilizer formalism must be represented by a contextual ontological model, when transformations are included in the operational description. This is contrary to the previously held belief that the single-qubit stabilizer formalism was a non-contextual subtheory of a single qubit.

Joel J. Wallman had the initial realization that the 8-state model of the stabilizer formalism is transformation contextual, and then gave the project to me to show that any ontological model of the single-qubit stabilizer formalism must be contextual, which eventually resulted in the paper given below. Joseph Emerson provided supervision and detailed revisions of the paper. The paper has been reproduced as it was published. However slight typographical changes have been made to fit with the thesis' overall style.

Additionally, I have included in this chapter some follow on work from the project. Presenting an alternative proof of transformation contextuality for the single-qubit stabilizer formalism, which uses the depolarizing channel and a coarse-grained measurement update to derive a proof of transformation contextuality. This proof has some appealing features when compared to the proof presented in the paper, and raises a few interesting questions. I have also included our investigation into how the results in the paper can be related to

proofs of contextuality in higher dimensions, via the Choi-Jamilkowski isomorphism. We were unable to find a satisfactory connection, but the work did produce some interesting results.

Contextuality and the Single-Qubit Stabilizer Subtheory

Piers Lillystone[1], Joel J. Wallman[2], and Joseph Emerson[2,3]

[1]*Institute for Quantum Computing and Department of Physics and Astronomy, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada*
[2]*Institute for Quantum Computing and Department of Applied Mathematics, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada*
[3] *Canadian Institute for Advanced Research, Toronto, Ontario M5G 1Z8, Canada*

Contextuality is a fundamental non-classical property of quantum theory, which has recently been proven to be a key resource for achieving quantum speed-ups in some leading models of quantum computation. However, which of the forms of contextuality, and how much thereof, are required to obtain a speed-up in an arbitrary model of quantum computation remains unclear. In this paper, we show that the relation between contextuality and a compuational advantage is more complicated than previously thought. We achieve this by proving that generalized contextuality is present even within the simplest subset of quantum operations, the so-called single-qubit stabilizer theory, which offers no computational advantage and was previously believed to be completely non-contextual. However, the contextuality of the single-qubit stabilizer theory can be confined to transformations. Therefore our result also demonstrates that the commonly considered prepare-and-measure scenarios (which ignore transformations) do not fully capture the contextuality of quantum theory.

1.2   Introduction

Contextuality [7, 43, 55, 63, 56, 76, 75, 20], which includes the better-known concept of Bell non-locality as a special case, is often regarded as the fundamental non-classical property of quantum theory. Furthermore, contextuality has emerged as an intriguing explanation for the power of quantum computation: as contextuality is required [39, 25, 9, 71] to achieve an exponential quantum speed-up by injecting magic states into Clifford circuits [15]; and also quantifies the computational advantage that can be obtained [82, 83, 9, 61, 21] in both the magic-state and measurement-based models of quantum computation [70].

Contextuality has also proved key in understanding quantum correlations, formalized in the Cabello-Severini-Winters graph formalism [20]. This formalism gives experimentally

testable traditional contextuality inequalities [17, 42, 85], parallelling non-locality inequalities such as CHSH inequality [22, 16]. Similar inequalities for generalized contextuality can also be constructed, and have already been experimentally verified [54]. Contextuality has also found applications in quantifying the memory cost of simulating quantum processes [40, 41, 19].

These considerations motivate us to understand the scope of phenomena that exhibit contextuality, with the aim of identifying which features of contextual phenomena enable quantum computational speed-up. However, one of the primary obstacles to understanding how contextuality powers a quantum computer is that the multi-qubit stabilizer subtheory [1] exhibits contextuality and yet can be efficiently simulated on a classical computer [31, 2].

There are two leading definitions of contextuality: traditional contextuality [7, 43, 55, 63, 56, 20], often referred to as Bell-Kochen-Specker contextuality, and generalized contextuality [76, 75]. In this paper, we show that, generalized contextuality [76] is present even in the single-qubit stabilizer subtheory of quantum theory, a fact missed by previous work [84, 11, 44]. We further demonstrate that the contextuality present in the single-qubit stabilizer subtheory can be confined to only appear in the transformations. This contradicts the common — and often implicit — assumption that an operational theory can be classified as contextual or non-contextual by only considering the preparations and measurements [35, 78, 46, 9, 45, 73].

1.3 Operational theories and ontological models thereof

An operational theory is non-contextual under a given definition if there exists an ontological model of the operational theory satisfying a specific property that we describe below.

An operational theory is a mathematical framework for predicting the outcomes of an experimental procedure, that is, a sequence of preparations, transformations, and measurements. These elements fully determine the experimental statistics, that is, the conditional probabilities $\Pr(k|P, T, M)$ of observing the outcome k when the preparation P, transformation T, and measurement M are performed sequentially. In quantum mechanics, the conditional probabilities for an experiment consisting of preparing a density matrix

---

[1]A subtheory of quantum theory is a closed subset of the operations available within quantum theory for a given Hilbert space dimension.

ρ, applying a completely positive and trace-preserving (CPTP) map Φ, and measuring a positive-operator-valued measure (POVM) $\{E_k\}$ are $\Pr(k|\rho, \Phi, E_k) = \text{Tr}(\text{E}_\text{k}\Phi(\rho))$ by the Born rule.

We can describe the underlying physical processes that generate the experimental statistics using the ontological models formalism. Here we follow the treatment in Ref. [50]. An ontological model is defined by a measurable space Λ of possible physical states, with an associated σ-algebra Σ, and sets of measures or measurable functions over Λ are used to represent preparations, transformations and measurements in the ontological model. For simplicity, we assume that there exists a measure that dominates all other measures in the model [2] (see the Appendix for a proof of the main theorem without this assumption). This allows us to express an ontological model in terms of probability densities, stochastic matrices, and response functions.

When a system is prepared via some procedure P, the physical properties of the system are probabilistically assigned values, which are completely encoded by the physical states $\lambda \in \Lambda$. Mathematically, we associate each preparation procedure P with a probability density over Λ, $\mu_P : \Lambda \to [0, 1]$, where $\int_\Lambda \mu_P(\lambda)d\lambda = 1$ as a system is always in some physical state. That is, the probability that a physical state, $\lambda \in \Lambda$, was prepared via P is $\mu_P(\lambda)$.

Similarly, when a transformation is applied to a system, the physical properties of the system dynamically evolve according to some stochastic map. Formally, we associate each transformation procedure T with a stochastic map $\Gamma_T : \Lambda \times \Lambda \to [0, 1]$, where the conditional probability that some λ is sent to another λ' by T is $\Gamma_T(\lambda', \lambda)$. As every physical state is mapped to some physical state by a transformation, $\int_\Lambda \Gamma_T(\lambda', \lambda)d\lambda' = 1$ for all $\lambda \in \Lambda$.

Finally, when a system is measured via some procedure M, the probability that outcome k occurs is specified by the physical state. That is, a measurement M is equivalent to a set of conditional probability functions $\{\xi_k^M : \Lambda \to [0, 1]\}_k$. As some measurement outcome always occurs, $\sum_k \xi_k^M(\lambda) = 1$ for all $\lambda \in \Lambda$. To correctly reproduce the experimental statistics of the operational theory, the distributions must satisfy

$$\Pr(k|P, T, M) = \int_\Lambda \xi_k^M(\lambda')\Gamma_T(\lambda', \lambda)\mu_P(\lambda)d\lambda d\lambda'. \tag{1.1}$$

Ontological models are assumed, often implicitly [12], to be convex-linear, that is, a probabilistic implementation of a set of operations is represented by the probabilistic mixture of the corresponding probability densities.

---

[2]That is $\exists \sigma$ such that $\sigma(\Delta) = 0 \Rightarrow \mu(\Delta) = 0$, $\Delta \in \Sigma$ for all μ in the model.

1.4 Generalized Contextuality

We now review generalized contextuality. The (experimental) setting of an operation is the set of other operations that are performed with the operation during an experiment. Two operations are operationally equivalent, denoted $\cong$, if they produce the same outcome statistics in all settings.

- Two preparations P and P' are equivalent, $(P \cong P')$,
  if $\Pr(k|P, T, M) = \Pr(k|P', T, M) \; \forall T, M$;

- Two transformations T and T' are equivalent,
  $(T \cong T')$, if $\Pr(k|P, T, M) = \Pr(k|P, T', M) \; \forall P, M$; and

- Two measurement outcomes $k \in M$ and $k \in M'$ are equivalent, $[(k, M) \cong (k, M')]$,
  if $\Pr(k|P, T, M) = \Pr(k|P, T, M') \; \forall P, T$.

Note that the definition of operational equivalence differs slightly from that of Ref. [76] in that we consider operational equivalence of individual measurement outcomes. However, this definition can be obtained from that of Ref. [76] by coarse-graining all measurements into two-outcome POVMs [49, 79, 46].

An ontological model is preparation non-contextual (PNC) if operationally equivalent preparation procedures are represented by the same probability densities, that is,

$$P \cong P' \Rightarrow \mu_P = \mu_{P'}. \tag{1.2}$$

Similarly, an ontological model is transformation non-contextual (TNC) if

$$T \cong T' \Rightarrow \Gamma_T = \Gamma_{T'} \tag{1.3}$$

and measurement non-contextual (MNC) if

$$(k, M) \cong (k, M') \Rightarrow \xi_{k,M} = \xi_{k,M'}. \tag{1.4}$$

An ontological model is universally non-contextual, in the generalized sense, if it satisfies eqs. (2.2) to (2.4), otherwise it is contextual [76]. Note, these definitions are trivially bi-directional, but to retain the definitional understanding we have left the implication from operationally equivalence to ontological equivalence.

Even a single qubit manifests generalized contextuality [76]. However, previous proofs of generalized contextuality for a single qubit have required subtheories strictly larger the single-qubit stabilizer subtheory.

Previous proofs of generalized contextuality have mostly been focused on the prepare-measure setting, as defined in [50], where in transformations are considered part of a preparation or measurement procedure, and systems are discarded post-measurement. However, this seemingly innocuous operational assumption is insufficient to identify an operational theory as non-contextual. As the contextuality of the operational theory *can* be confined to transformations, as we shall see is the case with the single-qubit stabilizer subtheory.

1.5 Contextuality in the 8-state model

We now show that the 8-state model of the single-qubit stabilizer subtheory exhibits transformation contextuality, a feature missed in previous studies of this model [84, 11, 44]. The single-qubit stabilizer subtheory consists of preparations and measurements in the eigen-bases of the single-qubit Pauli matrices $\{X, Y, Z\}$, the group of unitary transformations that permute the signed single-qubit Pauli matrices (i.e. the single-qubit Clifford group) and convex combinations of these operations. The single-qubit stabilizer subtheory has the property that preparing an eigenstate of one Pauli matrix P with eigenvalue η then measuring another Pauli Q results in the eigenvalue $\eta' = \pm\eta$ if $P = \pm Q$ and otherwise results in either eigenvalue with equal probability.

The 8-state model, originally developed in Ref. [84], is a natural ontological model for the single-qubit stabilizer subtheory (see fig. 1.1). It is defined by setting $\Lambda = \{\pm 1\}^{\times 3}$ and writing $\lambda = (x, y, z)$, where x, y, and z are the eigenvalues of X, Y, and Z respectively. These ontic states form the extremal points of the classical probability simplex for three random binary variables. Preparing the η eigenstate of X corresponds to setting $x = \eta$ and choosing y and z uniformly at random, *etc.* Similarly, measuring X returns the value of x, *etc.* This model is both preparation and measurement non-contextual [84].

In the 8-state model, a transformation corresponds to a permutation that acts on the hidden variable (x, y, z) in the same way that it acts on the Pauli operators (X, Y, Z). For example, conjugation by X maps $(X, Y, Z) \to (X, -Y, -Z)$, and so is represented by the permutation $\Gamma_X : (x, y, z) \to (x, -y, -z)$, with the transformations for Y and Z defined in a similar manner. Conjugation by the Hadamard matrix,

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},$$

maps $(X, Y, Z) \to (Z, -Y, X)$ and so is represented by the permutation $\Gamma_H : (x, y, z) \to (z, -y, x)$. Note that a Pauli operation preserves the parity xyz and the Hadamard swaps it (as does the phase gate P).
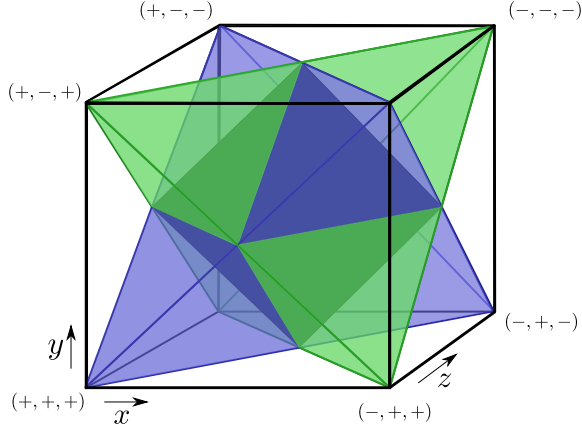
Figure 1.1: A graphical representation of the ontic space Λ of the 8-state model, where the tuples (x, y, z) index ontic states. The green and blue tetrahedra are the simplices of the odd- and even-parity ontic states respectively. The stabilizer polytope is the octahedron defined by the intersection of the two tetrahedra. The transformation $\Gamma_{\mathcal{T}_1}$ from eq. (1.6) maps any ontic state in one of the tetrahedra to another ontic state in the same tetrahedra, while the operationally equivalent transformation $\Gamma_{\mathcal{T}_2}$ from eq. (1.7) maps any ontic state to an ontic state in the opposite tetrahedra.

The 8-state model can be viewed as a conjunction of two distinct generalized non-contextual 4-state Wigner functions [77, 84, 9]. By extending the 4-state Wigner function to the 8-state model, all Clifford operations can be ontologically represented. As we now show, this increase in the size of the ontology causes the 8-state model to be transformation contextual.

To show the 8-state model is transformation contextual, let

$$\mathcal{T}_1(\rho) = [\rho + X\rho X + Y\rho Y + Z\rho Z]/4$$
$$, \mathcal{T}_2(\rho) = H\mathcal{T}_1(\rho)H. \tag{1.5}$$

These two transformations are operationally equivalent, as $\mathcal{T}_1(\rho) = \mathcal{T}_2(\rho) = I/2$ for any input state ρ. However, by convexity we have

$$\Gamma_{\mathcal{T}_1}[(a, b, c), (x, y, z)] = \begin{cases} \frac{1}{4} & \text{if } xyz = abc \\ 0 & \text{otherwise,} \end{cases} \tag{1.6}$$

18

while, since the Hadamard swaps the sign of xyz,

$$\Gamma_{\mathcal{T}_2}[(a,b,c),(x,y,z)] = \begin{cases} 0 & \text{if } xyz = abc \\ \frac{1}{4} & \text{otherwise.} \end{cases} \qquad (1.7)$$

That is, $\Gamma_{\mathcal{T}_1} \neq \Gamma_{\mathcal{T}_2}$, as illustrated in fig. 1.1.

1.6 The single-qubit stabilizer subtheory is contextual

Above we demonstrated that the 8-state model for the single-qubit stabilizer subtheory is transformation contextual. We now prove that there is no generalized non-contextual model for the single-qubit stabilizer subtheory, and hence that the single-qubit stabilizer subtheory is contextual. The proof follows by reducing the ontic space of a general preparation non-contextual model of the single-qubit stabilizer subtheory to that of the 8-state model.

> **Theorem 1.6.1** *Every ontological model of the single-qubit stabilizer subtheory is either preparation or transformation contextual.*

Proof. Fix an arbitrary preparation non-contextual ontological model of the single-qubit stabilizer subtheory. Let Δ_ρ be the support of the quantum state ρ in the ontological model, that is the set of physical states ρ has some possibility of preparing,

$$\Delta_\rho = \{\lambda | \, \mu_\rho(\lambda) > 0, \lambda \in \Lambda\}. \qquad (1.8)$$

Deleting any ontic state $\lambda \in \Lambda$ such that $\mu_{I/2}(\lambda) = 0$, we can partition Λ into 8 disjoint spanning sets from the assumption of PNC [76, eqs. (11) and (83)–(87)],

$$\Lambda_{x,y,z} = \Delta_{(I+xX)/2} \cap \Delta_{(I+yY)/2} \cap \Delta_{(I+zZ)/2}. \qquad (1.9)$$

As the model is preparation non-contextual, every quantum state has a unique support. Hence this partitioning is unique.

Noting that preparing σ and then applying a transformation T, that implements a CPTP map Φ, is a valid preparation procedure for the state $\Phi(\sigma)$. It must be the case that

19

T maps the support of ρ to the support of $\Phi(\rho)$ in a preparation non-contextual ontological model;

$$\Gamma_T : \Delta_\rho \to \Delta_{\Phi(\rho)}. \tag{1.10}$$

Therefore a Pauli X unitary must be represented by the permutation of the supports $\tau_X : \Lambda_{x,y,z} \to \Lambda_{x,-y,-z}$ on the partition $\{\Lambda_{x,y,z}\}$. Similarly, Pauli I, Y, and Z transformations must be represented by the respective permutations of the supports $\tau_I : \Lambda_{x,y,z} \to \Lambda_{x,y,z}$, $\tau_Y : \Lambda_{x,y,z} \to \Lambda_{-x,y,-z}$, and $\tau_Z : \Lambda_{x,y,z} \to \Lambda_{-x,-y,z}$. Therefore by convex linearity there exists an implementation of $\mathcal{T}_1$ that has the same stochastic map as eq. (1.6), when defined over the coarse-grained sets $\Lambda_{x,y,z}$.

Similarly for the Hadamard gate, we have the map $\tau_H : \Lambda_{x,y,z} \to \Lambda_{x,-y,z}$. Therefore there exists an implementation of $\mathcal{T}_2$ that has the same stochastic map as eq. (1.7), when defined over the coarse-grained sets $\Lambda_{x,y,z}$. That is $\mathcal{T}_1 \cong \mathcal{T}_2$ and yet they cannot be represented by the same stochastic map in any preparation non-contextual model.

∎

We now show that any model of the single-qubit stabilizer subtheory must be either traditionally contextual [43] or transformation contextual. We do this by proving a stronger result;

> **Theorem 1.6.2** *Every ontological model of the single-qubit stabilizer subtheory is either outcome-indeterministic or transformation contextual.*

Proof. The proof proceeds in the same manner as theorem 1.6.1, where now we assume outcome determinism rather than preparation non-contextuality. By outcome determinism, we can partition Λ into 8 disjoint sets according to measurement outcomes;

$$\widetilde{\Lambda}^{X,Y,Z}_{xyz} = \left\{ \lambda | \xi^X_x(\lambda) = 1, \, \xi^Y_y(\lambda) = 1, \, \xi^Z_z(\lambda) = 1 \right\}, \tag{1.11}$$

where our choice of measurement contexts X, Y, and Z is arbitrary and is merely used to partition Λ such that the maps representing Clifford transformations are well defined. Using the equivalent to eq. (1.10) for measurements, i.e. how transformations map the supports defined by eq. (1.11), the maps $\Gamma_{\mathcal{T}_1}$ and $\Gamma_{\mathcal{T}_2}$ must be represented as stated in eqs. (1.6) and (1.7). ∎

Therefore theorem 1.6.2 implies any traditionally non-contextual ontological model of the single-qubit stabilizer formalism, for example Kochen-Specker's original model of a single qubit [43], must be transformation contextual, as traditional contextuality is implied by the conjunction of outcome determinism and generalized measurement non-contextuality [76]. Finally we note theorem 1.6.2 can be used to prove theorem 1.6.1, as preparation non-contextuality implies traditional non-contextuality for sharp measurements [48].

1.7 Discussion

In this Letter we have shown that the single-qubit stabilizer subtheory, a very simple subtheory of the smallest quantum system, exhibits generalized contextuality. This demonstrates that generalized contextuality is so prevalent that even an essentially trivial quantum subtheory is classified as contextual, and therefore non-classical. The result shows that, unlike traditional contextuality which may be an important resource for quantum speed-up, generalized contextuality is an extremely weak notion of contextuality, that exists in trivial and classical-like models, such as the 8-state model. This may reflect the fact that generalized contextuality is defined in terms of a context-dependence at the epistemic level, rather than a context-dependence at the ontological level.

The contextuality in the single-qubit stabilizer subtheory is only apparent if all operations are accounted for, that is all stabilizer states, all stabilizer measurements, and the full Clifford group. Therefore a universally non-contextual model can only be constructed for strict subtheories of the singe-qubit stabilizer subtheory. For example, the Hadamard and Phase gates are not elements of the toy theory [77] or the standard Wigner function [29], conversely the Hadamard gate is an element of the rebit subtheory [25], but Y eigenstates and Y measurements are not.

Our result also demonstrates that the operational reduction of only considering preparations and measurements is less robust than previously recognized. As this reduction can conceal key features of the model, such as the presence of some forms of contextuality. It is an interesting open problem to understand how and when this kind of reduction can obscure such important conceptual features of an operational theory. This insight regarding the key role of transformations may have unexplored connections to research into the memory cost of quantum simulations [40, 41, 19], and the formalism of computational mechanics [5, 80].

Another possible route to investigate the role of transformations in physical theories is

the Choi-Jamiołkowski isomorphism. The isomorphism, in quantum theory, relates transformations to states in a larger Hilbert space. Hence by using a similar isomorphism for ontological models we may be able to find a connection between the impossibility of a universal non-contextual ontological model of the single qubit stabilizer subtheory and the impossibility of a preparation non-contextual ontological model of the two-qubit stabilizer subtheory [56, 48].

1.7.1 Acknowledgments

1.8 Appendix: Measure-Theoretic Treatment of Theorem 1

We now prove theorem 1.6.1 in the more general measure-theoretic framework for ontological models, see [50] definition 8.2.

> **Theorem 1.8.1** *Every ontological model of the single-qubit stabilizer subtheory is either preparation or transformation contextual.*

Proof. To adapt the proof to the measure-theoretic setting, we need to change the definition of support of a quantum state as follows. Let Δ_ρ be the support of the quantum state ρ in the ontological model, that is, a (not necessarily unique) set such that for all $S \in \Sigma$,

$$\mu_\rho(S) \begin{cases} = 1 & \text{if } \Delta_\rho \subseteq S \\ < 1 & \text{otherwise.} \end{cases} \tag{1.12}$$

As before we delete any measurable set $S \in \Sigma$ such that $\mu_{I/2}(S) = 0$, then partition Λ into 8 spanning sets that intersect on sets of measure zero, from the assumption of PNC [76, eqs. (11) and (83)–(87)]. Having reduced the model to a model over a finite set of states, the rest of the proof follows as described in the main text. ∎

1.9 Supplementary Material

1.9.1 A Measurement Based Proof of Transformation Contextuality

The proof of transformation contextuality given in this paper is only one example of a set of proofs of transformation contextuality for the single-qubit stabilizer subtheory. Clearly the transformation $\mathcal{T}_2$ can have the Hadamard replaced with another Clifford operation. To be exact there are 8 choices which can be constructed by replacing H with any element of the set $\{H, S\} \circ \{\mathbb{I}, X, Y, Z\}$. Additionally we can replace $\mathcal{T}_2$ with the channel;

$$\mathcal{T}_C(\rho) = \frac{1}{24} \sum_{C \in Cl_1} C\rho C^\dagger,$$

which results in $\Gamma_{\mathcal{T}_C}$ uniformly randomizing the input ontic state over the full ontology. There are almost certainly more proofs of transformation contextuality for the single-qubit stabilizer formalism that are based on convex combinations of Clifford channels, beyond those given above. However, in this section I shall present a more exotic proof of transformation contextuality that begins to link transformation contextuality to measurements. This style of proof also has a few interesting features that distinguish it from the Clifford based proofs.

The maps that we use in this proof of contextuality are the depolarizing channel, for a given Pauli, and a coarse-graining over a measurement in the respective basis, i.e.;

$$\mathcal{T}_{D_P}(\rho) = \frac{1}{2} \left(\rho + P\rho P \right), \tag{1.13}$$

$$\mathcal{T}_{M_P}(\rho) = \Pi_{+1,P}\rho\Pi_{-1,P} + \Pi_{-1,P}\rho\Pi_{-1,P}. \tag{1.14}$$

By expanding $\Pi_{\pm 1,P} = (\mathbb{I} \pm P)/2$ we can easily verify that these two channels are operationally equivalent for all ρ.

The ontological map for the Pauli operation, Γ_P, behaves as described in the main body of the paper. The measurement update map for a measurement of P with outcome k, $\Gamma_{k,P}$, maps an ontic state (x, y, z) to an ontic state where the bit corresponding to the measurement is the same as prior to the measurement, for repeatability, and uniformly randomizes the other bits, to destroy information about the state prior to measurement. From this we can see that $\mathcal{T}_{D_P}$ and $\mathcal{T}_{M_P}$ have different representations and therefore are contextually represented in all ontological models of the single-qubit stabilizer formalism[3].

---

[3]By the same arguments as presented in the paper.

To be more explicit consider a depolarizing map and corresponding measurement followed by a coarse-graining in the Z-basis;

$$\mathcal{T}_{D_Z}(\rho) = \frac{1}{2}\left(\rho + Z\rho Z\right),$$

$$\mathcal{T}_{M_Z}(\rho) = \Pi_{+1,Z}\rho\Pi_{-1,Z} + \Pi_{-1,Z}\rho\Pi_{-1,Z}.$$

These two CPTP maps have ontological maps given by;

$$\Gamma_{\mathcal{T}_{D_Z}} : (x, y, z) \mapsto (x + a, y + a, z),$$

with $a \in \{0, 1\}$ with equal probability, and

$$\Gamma_{\mathcal{T}_{M_Z}} : (x, y, z) \mapsto (x + a, y + b, z)$$

with $a, b \in \{0, 1\}$ with equal probability. Hence, as $\mathcal{T}_{D_Z} \cong \mathcal{T}_{M_Z}$ but $\Gamma_{\mathcal{T}_{D_Z}} \neq \Gamma_{\mathcal{T}_{M_Z}}$ these maps are contextually represented.

So what distinguishes this style of proof from the one given in the paper?

1. The inclusion of measurements in the map would hopefully convince some of the more die-hard operationalist that there is some merit in investigating transformations[4].

2. The proof seemingly does not need the full Clifford group. However sadly this is not the case, Clifford operations are required for this proof to hold. The assumption that the ontology can be partitioned into the 8 ontic regions, indexed by the tuple (x, y, z), can be derived by considering the action of Clifford operations on the representation of stabilizer states. For a counter-example consider that Spekkens' toy-model [77] non-contextually represents $\Gamma_{\mathcal{T}_{D_Z}}$ and $\Gamma_{\mathcal{T}_{M_Z}}$, due to the lack of the y-bit to randomize.

3. This proof only holds for the full single-qubit stabilizer subtheory. Both Spekkens' toy-model and the single-rebit subtheory, we shall see later that the single-rebit subtheory requires a contextual ontological model, represent these transformations non-contextually.

---

[4]Given one of the main qualms about the paper is statements to the effect *"There are no transformations, just preparations and measurements"*, which personally I feel is far too philosophically motivated, even for a very philosophical field of physics. To combat this viewpoint I suggest an *all-or-nothing* doctrine for the ontological models formalism. That is to say, an ontological model should reproduce *all* the physical features of an operational theory to be considered a physical account of said theory, not just some subset of operational elements. Non-destructive measurements exist and we can't ignore them.

4. Unlike the transformations used in the paper, where $\mathcal{T}_1(\rho) = \mathcal{T}_2(\rho) = \mathbb{I}/2$, $\forall \rho$, the CPTP maps $\mathcal{T}_{D_z}(\rho)$ and $\mathcal{T}_{M_z}(\rho)$ do *not* map all states to the maximally mixed state, making them operationally more interesting.

5. If we look at the 8-state Wigner function we find that the representation of the measurement update maps $\Gamma_{k,P}$ are positively represented. Unlike the Clifford operations, for which all Clifford operations are negatively represented including the identity. Therefore a map being positive does not preclude its use in a proof of contextuality.

This proof raises one important question in my eyes, is the definition of measurement non-contextuality complete? Equation (2.4) is the standard definition of measurement non-contextuality and is certainly an adequate one if we only consider destructive measurements. However, if we allow state updates post measurement, i.e. non-destructive measurements, then this definition is missing the critical component of the measurement-update map. When these update maps are included in an operational description of a theory many of the operational equivalences disappear. To see this consider that the Kraus-operators associated to an update after a measurement of the POVM element E_k only need to satisfy $E_k = M_k^\dagger M_k$, which allows us to distinguish between measurements of E_k in different "contexts", by using sequential measurements on the system.

This leaves us with two possibilities for extending the definition of measurement non-contextuality. In the first we keep the definition of MNC and add a fourth condition *measurement-update* non-contextuality (MUNC). I.e. An ontological model is measurement update non-contextual if all operationally equivalent measurement updates are represented by the same probability densities;

$$\Gamma_{k,M} = \Gamma_{k,M'} \Leftrightarrow T_{(k,M)} \cong T_{(k,M')}. \tag{1.15}$$

This definition then implies measurement non-contextuality[5].

The second, and more lenient, definition is to demand that the joint distribution composed of the response function and measurement update map, for a given measurement outcome, are equal for all operationally equivalent measurements. Let us call this definition of non-contextuality *complete measurement non-contextuality* (CMNC) that is;

$$\Gamma_{k,M}\xi_{k,M} = \Gamma_{k,M'}\xi_{k,M'} \Leftrightarrow (k, M, T_{(k,M)}) \cong (k, M, T_{(k,M')}). \tag{1.16}$$

---

[5] As pointed out by R. Spekkens during the defense of this thesis.

In my opinion both of these definitions present very interesting research avenues, because they are far more aligned with the original controversy surrounding measurements in quantum theory, such as the Einstein-Podolsky-Rosen paradox and subsequently Bell's theorem. Reinstating the role of measurement update in foundational discussions.

1.9.2 Connecting Transformation Contextuality to Traditional Contextuality

One of the open questions raised by the paper presented in this chapter is how does this new proof of transformation contextuality relate to older proofs of contextuality, particularly traditional contextuality. To this end, Joel J. Wallman suggested we look at the Choi–Jamiołkowski (CJ) isomorphism from quantum theory [60]. Hence, prior to moving on to other projects, I investigated whether these two types of contextuality could be related in a satisfactory way. Unfortunately, we were unable to find a rigorous relation between the two, but did find a weaker connection and were able to make some observations on state-dependent proofs of contextuality and the stabilizer formalism. In this section, we shall explore the possible links between transformation and traditional contextuality, in the context of the paper. However, in the interest of space, we will not dive too deep into the background, but rather gives references where necessary and more focus on the big picture.

The Choi-Jamiołkowski (CJ) isomorphism [60] allows us to connect completely-positive (CP) maps on $\mathcal{H}_d$ to quantum states in $\mathcal{H}_{2d}$. The standard definition of the CJ isomorphism associates a CP map $\mathcal{T} \in \mathcal{L}(\mathcal{H}_d)$ to some state $\rho_{\mathcal{E}} \in D(\mathcal{H}_{2d})$ via;

$$\rho_{\mathcal{E}} = (\mathcal{T} \otimes \mathbb{I}) (|B_{00}\rangle \langle B_{00}|). \tag{1.17}$$

The first thing to note is that the transformations used in the proof of transformation contextuality, $\mathcal{T}_1$ and $\mathcal{T}_2$, both have trivial representations under the CJ isomorphism. I.e. $\rho_{\mathcal{T}_1} = \rho_{\mathcal{T}_2} = \mathbb{I}/4$. Therefore, we will focus our attention on the single-qubit Clifford gates.

The single-qubit Clifford gates, e.g. $C \in Cl_2$, under the CJ ismorphism are mapped to the two-qubit entangled stabilizer states, $\left\{ \Pi_{ent}^{(C)} \right\}$;

$$\Pi_{ent}^{(c)} = (\mathcal{T}_C \otimes \mathbb{I}) (|B_{00}\rangle \langle B_{00}|),$$
$$= (C \otimes \mathbb{I}) |B_{00}\rangle \langle B_{00}| (C \otimes \mathbb{I})^{\dagger} = \left| \psi_{ent}^{(C)} \right\rangle \left\langle \psi_{ent}^{(C)} \right|,$$

Where $|B_{00}\rangle = |00\rangle + |11\rangle$ is the unnormalized Bell state and local operations preserve entanglement. This gives a one-to-one correspondence between the 24 single qubit Clifford gates and the 24 two-qubit entangled stabilizer states.

To attempt to connect the proof of transformation contextuality and traditional contextuality we first note that the transformations used in the proof are of the form $\mathcal{T}_C(\rho) = C\mathcal{T}_D(\rho)C^\dagger$, where we have relabelled $\mathcal{T}_1 \to \mathcal{T}_D$[6] such that $\mathcal{T}_1 = \mathcal{T}_\mathbb{I}$ and $\mathcal{T}_2 = \mathcal{T}_H$. Therefore, by finding a connection between the two sets $\left\{\Pi_{ent}^{(C)}\right\}$ and $\{\mathcal{T}_C\}$ we may be able to understand how transformation contextuality for a single-qubit and traditional contextuality for two-qubits are related.

To investigate contextuality proofs using states from $\left\{\Pi_{ent}^{(C)}\right\}$ we will use the Cabello-Severini-Winter (CSW) exclusivity graph formalism [20]. As a brief overview an exclusivity graph is constructed by associating projectors to vertices of a graph and edges are drawn between commuting projectors. From this graph we can then investigate a variety of properties that have a one-to-one correspondence with contextuality inequalities. We can also look at subgraphs of any exclusivity graph to find additional proofs of contextuality [18].

Using the CSW exclusivity graph formalism we can construct the exclusivity graph, G_{ent}, for the entangled set of 2-qubit stabilizer states, $\left\{\Pi_{ent}^{(C)}\right\}$, given in fig. 1.2. This exclusivity graph is equivalent to the exclusivity graph that would be constructed for the Mermin-Peres square given in fig. 1.3. Therefore fig. 1.2 constitutes a state-independent proof of traditional contextuality. By numerical calculation it is straightforward to verify that the graph properties, relevant to contextuality, are given by;

$$\text{Indepence Number:} \quad \alpha(G_{ent}) = 5,$$
$$\text{Theta value:} \quad \theta(G_{ent}) = 6,$$
$$\text{Fractional Packing:} \quad \alpha^\star(G_{ent}) = 6.$$

Before we start to connect transformation contextuality to traditional contextuality, we first need to define the notion of an entangled stabilizer state's *phase class*. An entangled stabilizer state can be expressed by its stabilizer group as follows;

$$S(\Pi_{ent}) = \{(-1)^{p_0}\mathbb{I}, (-1)^{p_1}P_1, (-1)^{p_2}P_2, (-1)^{p_3}P_3|$$
$$P_i \in \{\{X, Y, Z\} \otimes \{X, Y, Z\}\}, \, p_{i\neq0} \in \mathbb{Z}_2, \, p_0 = 0\}, \quad (1.18)$$

where this is a group because $(-1)^{p_1}(-1)^{p_2}P_1P_2 = (-1)^{p_3}P_3$ and furthermore is Abelian as $[P_i, P_j] = 0, \, \forall i, j$. For a more detailed review of the n-qubit stabilizer formalism refer to

---

[6]I.e. $\mathcal{T}_1$ is the maximally depolarizing channel.
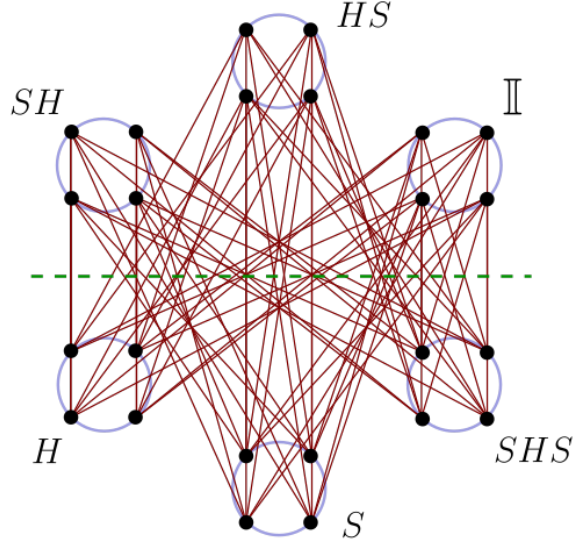
Figure 1.2: The exclusivity graph, G_{ent}, for the set $\left\{\Pi_{ent}^{(C)} \,|\, C \in Cl_{2^n}\right\}$. Each vertex represents a projector onto one of the entangled stabilizer states, organised as follows; The Bell-basis is the 4-vertices, connected by a blue circle, beside $\mathbb{I}$. The Clifford operation beside each blue circle, with 4 vertices, denotes which Clifford maps the Bell-basis to the associated basis, via applying the Clifford to the first qubit. The green dotted line then distinguishes the different phase classes of $\{\Pi_{ent}\}$, with the 3 bases above the line having unbalanced phase vectors and the ones below having balanced phase vectors. Red and blue lines connecting vertices indicate that the two projectors associated to these vertices commute.

$$
\begin{array}{ccccccc}
ZY & - & YZ & - & XX & \rightarrow & \mathbb{I} \\
| & & | & & | & & \\
XZ & - & ZX & - & YY & \rightarrow & \mathbb{I} \\
| & & | & & | & & \\
YX & - & XY & - & ZZ & \rightarrow & \mathbb{I} \\
\downarrow & & \downarrow & & \downarrow & & \\
-\mathbb{I} & & -\mathbb{I} & & -\mathbb{I} & &
\end{array}
$$

Figure 1.3: The Mermin-Peres square where-in all contexts have entangled simultaneous eigenbasis. Each context is defined by a row or column in the square, within which all observables commute, and tensor product notation has been suppressed.

28

Part III of this thesis. From eq. (1.18) we can define a 2-qubit entangled stabilizer state's phase vector by;

$$\vec{p}_{ent} = (p_0, p_1, p_2, p_3). \tag{1.19}$$

We then can partition the set of 2-qubit entangled stabilizer states into two equivalency classes $\mathcal{V}_1$ and $\mathcal{V}_2$, such that $\mathcal{V}_1 \cup \mathcal{V}_2 = \left\{ \Pi_{ent}^{(C)} \right\}$, by saying two states $\Pi_i \in \left\{ \Pi_{ent}^{(C)} \right\}$ and $\Pi_j \in \left\{ \Pi_{ent}^{(C)} \right\}$ are in the same phase equivalence class if their phase vectors are equivalent up-to the addition of some balanced vector, which can be defined by the inner product $\vec{v}_{\mathrm{Bal}} = [c, \cdot]$, i.e. a balanced vector where the zeroth entry is zero. So mathematically;

$$\vec{p}_i = \vec{p}_j + [c, \cdot] \iff \Pi_i, \Pi_j \in \mathcal{V}_k.$$

Returning to fig. 1.2 I have organised the two equivalences classes to be separated by the green dotted line, i.e. one is above the line and one is below the line. One of the powerful features of the CSW exclusivity graph formalism is that a contextuality proof can be constructed from a graph *iff* it contains odd-cycle subgraphs [18][7], of which fig. 1.2 contains many. With this in mind, fig. 1.2 clearly shows that to construct a proof of traditional contextuality with projectors from $\left\{ \Pi_{ent}^{(C)} \right\}$ we must use projectors from both $\mathcal{V}_1$ and $\mathcal{V}_2$.

Considering transformation contextuality, let us construct a similar partitioning of the set of transformations $\left\{ \mathcal{T}_C = C\mathcal{T}_D C^\dagger \right\}_C$. Namely, we shall partition them according to whether they are represented by maps of the form eq. (1.6) or eq. (1.7), let us call each partition $\mathcal{W}_1$ and $\mathcal{W}_2$. Therefore a proof of transformation contextuality can only be constructed by using a map from $\mathcal{W}_1$ *and* a map from $\mathcal{W}_2$. Explicitly these partitions are given by;

$$\mathcal{W}_1 = \left\{ \mathcal{T}_C | \, C \in \{\mathbb{I}, HS, SH\} \circ \{\mathbb{I}, X, Y, Z\} \right\}, \tag{1.20}$$
$$\mathcal{W}_2 = \left\{ \mathcal{T}_C | \, C \in \{H, S, SHS\} \circ \{\mathbb{I}, X, Y, Z\} \right\}, \tag{1.21}$$

where H is the Hadamard gate, and S is the phase gate.

With these partitions set up we can begin to qualitatively connect transformation contextuality and traditional contextuality. Namely, that the CJ representation of the Clifford gates used in $\mathcal{W}_1$ are exactly the states in $\mathcal{V}_1$, and similarly for $\mathcal{W}_2$ and $\mathcal{V}_2$. While this is not a quantitative or rigorous connection between the two, I believe it could lead to a

---

[7]Of size greater than 5.

more concrete link between the two forms of contextuality. Possibly by fine graining the maps used in $\mathcal{T}_C$, i.e. decomposing them into their constituent Clifford channels, a more rigorous connection can be found. Additionally, another relationship can be drawn between these sets; Any Clifford operation defining the transformations in the set $\mathcal{W}_1$ leaves the sets $\mathcal{V}_1$ and $\mathcal{V}_2$ invariant, i.e. $C\mathcal{V}_iC^\dagger = \mathcal{V}_i$, $\forall C \in \{\mathbb{I}, HS, SH\} \circ \{\mathbb{I}, X, Y, Z\}$. And any Clifford operation defining a map in $\mathcal{W}_2$ maps states in $\mathcal{V}_1$ to states in $\mathcal{V}_2$ and visa versa, i.e. $C\mathcal{V}_iC^\dagger = \mathcal{V}_{j\neq i}$, $\forall C \in \{H, S, SHS\} \circ \{\mathbb{I}, X, Y, Z\}$.

Alternatively, to find a connection between transformation and traditional contextuality one could investigate how the assumptions used in the proof of transformation contextuality can be "run-through" the CJ isomorphism. However, this mapping has some ambiguities, namely, there are many possibilities when extending a model from one-qubit to two-qubits. Further, the CJ isomorphism for channels seems to indicate an assumption of locality must be made. To see this, consider that one way of stating the CJ isomorphism is as a map from an input-output pair to an input-input pair, i.e. the CJ isomorphism performs a mapping of the form $|i\rangle\langle j| \rightarrow |i\rangle|j\rangle$. Does this imply that for the ontic input-output pair $\lambda_{in} \rightarrow \lambda_{out}$ we must use a ontology of the form $\lambda_{in} \times \lambda_{out} \in \Lambda \times \Lambda$?

5-cycles of the $\{\Pi_{ent}\}$ Exclusivity Graph

As a final aside, I'd like to note a kind of interesting property of certain 5-cycles that are subgraphs of fig. 1.2. The first thing to consider is that fig. 1.2 contains a very large number of subgraphs that contain odd-cycle subgraphs, and therefore any of these subgraphs provide a proof of contextuality. If we restrict our attention to subgraphs that are 5-cycles we find there is an interesting feature regarding the classical non-contextuality bound and the stabilizer formalism. Namely, the stabilizer formalism *cannot* saturate the classical non-contextuality inequality for some 5-cycles.

In one of the 5-cycles I investigated, the classical non-contextuality bound was $\alpha(G_5) = 2$, the quantum bound was $\theta(G_5) = 2.174...$, but the maximum value achievable via 2-qubit stabilizer states was $3/2$. Does this indicate that while the stabilizer subtheory requires a contextual ontological model, this contextuality is in some senses less powerful than a classical non-contextual model? Clearly this question is ill-posed at the moment, as we have very few if any satisfactory methods of quantifying how contextual a subtheory is, but with some refinement could lead to some interesting results. My conjecture is that this is related to Clifford circuits only being able to perform parity-L computation.

Chapter 2

Contextuality and Negativity for Subtheories of a Single Qubit

2.1 Chapter Preamble

In this chapter, I present work that was completed as another follow on project from the *Contextuality and the Single-Qubit Stabilizer Formalism* paper. The previous work had left a question open of how ubiquitous contextuality was once transformations were included in the operational description of a subtheory. It had also re-opened the question of how the necessity of a contextual ontological model was related to the necessity of a negatively represented quasi-probability representation, again once transformations were included in the operational description.

In this paper, to appear on the arXiv shortly, we answer the first question by showing that an even smaller subtheory of a qubit, the rebit subtheory, also requires a contextual description. Therefore implying that contextuality is far more prevalent than previous thought. We answer the second question by showing that both the single-qubit stabilizer subtheory and a singe-rebit subtheory require a negatively represented quasi-probability representation. Re-establishing the link between contextuality and negativity when transformations are included in the operational description.

This work presented here was completed in most part by me, Piers Lillystone. Joel J. Wallman helped streamline the transformations used in the proof of contextuality for the rebit subtheory and Joseph Emerson provided supervision over the project.

Contextuality and Negativity for Subtheories of a Single Qubit

Piers Lillystone[1], Joel J. Wallman[2], and Joseph Emerson[2,3]

[1]*Institute for Quantum Computing and Department of Physics and Astronomy, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada*
[2]*Institute for Quantum Computing and Department of Applied Mathematics, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada*
[3] *Canadian Institute for Advanced Research, Toronto, Ontario M5G 1Z8, Canada*

There has been a long established equivalence between the requirement of a contextual ontological model of quantum theory and quantum theory requiring a negatively represented quasi-probability representation. However, this equivalence was only explored in the context of prepare-measure scenarios. In this paper, we investigate the relationship between contextuality and negativity when transformations are included in the operational theory. Specifically, we show that the single-qubit stabilizer subtheory, which must admit a contextual ontological model, must be negatively represented by any quasi-probability representation. Further, we show that the single-rebit stabilizer subtheory must be represented by a contextual ontological model and be negatively represented by any quasi-probability representation. Contrary to the previously held belief that it is a non-contextual subtheory of the *n*-qubit stabilizer subtheory.

2.2 Introduction

Contextuality [7, 43, 55, 63, 56, 76, 75, 20] is a fundamental feature of quantum theory, which includes Bell non-locality as a special case, and has been proposed as *the* non-classical property of the theory that explains the dividing line between quantum and classical phenomena. Recently, contextuality has appeared as an intriguing explanation for the resource that powers quantum computation [39, 25, 9, 71]. Contextuality has been shown to be necessary for exponential speed-up in magic state injection schemes [15]. It can also be used to quantify computational advantages [82, 83, 9, 61, 21] in magic state and measurement based [70] models of quantum computation. Additionally, contextuality provides a tool for understanding the set of quantum correlations [20], and can be used to give experimentally testable inequalities [17, 42, 85, 74], paralleling the more familiar non-locality inequalities [22, 16]. More recently, contextuality has also been successful in lower bounding the memory cost of simulating quantum circuits [40, 41, 19].

The success of applying contextuality research to quantum information processing motivates us to investigate the scope of contextual phenomena. Recently, we showed that single-qubit stabilizer subtheory must admit a contextual ontological model [52], when transformations are included in the operational description of the subtheory. In this follow on paper, we show that the rebit subtheories [25], which are strict subtheories of the single-qubit stabilizer subtheory, do not also admit a non-contextual ontological model, when transformations are included in the operational description of these subtheories.

The requirement of a contextual ontological model of the rebit subtheory is somewhat surprising, given that the n-rebit subtheory admits a Kochen-Specker non-contextual ontological model [43, 25]. Further, the proof of contextuality used is novel in the sense that we utilize two completely-positive trace-preserving maps that are operationally equivalent within the subtheory, but not for larger theories, which includes the single-qubit stabilizer subtheory. Additionally, these operations only require Pauli unitaries, rather than relying on the specific structure of the allowed Clifford unitaries.

The necessity of a negative quasi-probability representation was previously shown to be equivalent to the necessity of a contextual ontological model [78, 28, 27]. However, this equivalence was only demonstrated in the prepare-measure experimental setting. In this paper we show that the single-qubit stabilizer subtheory and single-rebit stabilizer subtheory require a negative quasi-probability representation. Confirming the link between negativity and contextuality even when transformations are included in the operational theory.

2.3 Preliminaries

2.3.1 Ontological Models and Generalized Contextuality

An operational theory [36] is a methodology that mathematically captures the statistics of sets of experiments. A single experiment is considered to be composed of a sequence of preparations, transformations, and measurements. The sequence of operations performed then uniquely determines the outcome statistics of a given experiment. We express these outcome statistics with the conditional probability $\Pr(k|P, T, M)$, where k is the observed outcome of the performed measurement M, P is the preparation performed, and T is the transformation applied to the preparation.

The ontological models formalism aims to explain an operational theory's statistics by supposing there exists a physical model of an experiment's operational procedures. Here

we follow the definitions of an ontological model given in Ref. [50], but restrict ourselves to ontological models where there exists a measure that dominates all other measures[1].

The most basic assumption of the ontological models formalism is that there exist a set of real physical states that a system may be prepared in, which encode the physical properties of the system. We call these physical states *ontic* states, λ, and the set of all physical states the *ontology*, Λ. Preparation procedures are then considered to probabilistically prepare one of these physical states. Therefore, each preparation P is associated to a probability density, μ_P, over the ontology such that $\mu_P : \Lambda \mapsto [0, 1]$, i.e. μ_P gives the conditional probability of a preparation preparing a given ontic state, $\Pr(\lambda|P) = \mu_P(\lambda)$. As some physical state must be prepared we have $\int_\Lambda d\lambda \mu_P(\lambda) = 1$. Additionally, we define the support of a preparation procedure P to be $\Delta_P = \{\lambda|\, \mu_P(\lambda) > 0\}$.

Transformations are considered to stochastically evolve the system's ontic state. Therefore, each transformation T is associated to a stochastic map Γ_T that probabilistically maps the ontic states to ontic states such that $\Gamma_T : \Lambda \times \Lambda \mapsto [0, 1]$, i.e. Γ_T gives the conditional probability that T maps some ontic state λ to a new ontic state λ', $\Pr(\lambda'|\lambda, T) = \Gamma_T(\lambda', \lambda)$. As a transformation must map a physical state to another physical state we have $\int_\Lambda d\lambda' \Gamma_T(\lambda', \lambda) = 1$, $\forall \lambda \in \Lambda$.

Measurements in the ontological models formalism probabilistically output a possible outcome of the measurement. Therefore, for every measurement procedure M we associate a set of response functions $\{\xi_k^M : \Lambda \mapsto [0, 1]\}_{k \in M}$, with a response function defined for each outcome k, i.e. ξ_k^M gives the probability that outcome k was observed given a measurement M and some ontic state λ, $\Pr(k|\lambda, M) = \xi_k^M(\lambda)$. As a measurement outcome must always occur for any ontic state λ we have $\sum_{k \in M} \xi_k^M(\lambda) = 1$, $\forall \lambda \in \Lambda$.

An ontological model is said to correctly reproduce the statistics of an operational theory $\Pr(k|P, T, M)$ if we have;

$$\Pr(k|P, T, M) = \int_\Lambda d\lambda d\lambda' \xi_k^M(\lambda') \Gamma_T(\lambda', \lambda) \mu_P(\lambda), \ \forall P, T, (k, M). \tag{2.1}$$

2.3.2 Generalized Contextuality

Generalized contextuality [76] is defined via the notion of *operational equivalence*. Two operations are considered to be operationally equivalent, denoted $\cong$, if for all possible experimental settings the two operations produce the same experimental statistics, where an experimental setting is the other operational elements an operation is performed with.

---

[1]Allowing us to use probability densities rather than a full measure theoretic treatment.

Therefore for an operational theory, composed of a set of preparations $\mathcal{P}$, a set of transformations $\mathcal{T}$, and a set of measurements $\mathcal{M}$, we say: Two preparations $P \in \mathcal{P}$ and $P' \in \mathcal{P}$ are operationally equivalent, $(P \cong P')$, if $\Pr(k|P, T, M) = \Pr(k|P', T, M)$, $\forall T \in \mathcal{T}$, $\forall M \in \mathcal{M}$. Two transformations $T \in \mathcal{T}$ and $T' \in \mathcal{T}$ are operationally equivalent, $(T \cong T')$, if $\Pr(k|P, T, M) = \Pr(k|P, T', M)$, $\forall P \in \mathcal{P}$, $\forall M \in \mathcal{M}$. Two measurement outcomes $k \in M \in \mathcal{M}$ and $k \in M' \in \mathcal{M}$ are operationally equivalent, $[(k, M) \cong (k, M')]$, if $\Pr(k|P, T, M) = \Pr(k|P, T, M')$, $\forall P \in \mathcal{P}$, $\forall T \in \mathcal{T}$.

An ontological model is then considered non-contextual if operationally equivalent elements are represented by the same object in the ontological model. So, an ontological model is preparation non-contextual (PNC) if for all operationally equivalent preparations it satisfies;

$$P \cong P' \Rightarrow \mu_P = \mu_{P'}. \tag{2.2}$$

An ontological model is transformation non-contextual (TNC) if for all operationally equivalent transformations it satisfies;

$$T \cong T' \Rightarrow \Gamma_T = \Gamma_{T'}. \tag{2.3}$$

An ontological model is measurement non-contextual (MNC) if for all operationally equivalent measurement outcomes it satisfies;

$$(k, M) \cong (k, M') \Rightarrow \xi_{k,M} = \xi_{k,M'}.. \tag{2.4}$$

Therefore we say an operational theory is universally non-contextual, under the generalized notion of contextuality, if it admits a PNC (eq. (2.2)), TNC (eq. (2.3)), and MNC (eq. (2.4)) ontological model. Otherwise we say the operational theory admits only contextual ontological models [76].

2.3.3 Single-Qubit Stabilizer Subtheory

The single-qubit stabilizer subtheory is defined by the single qubit Pauli operators, mathematically $\mathcal{P}_1 = \{X, Y, Z\}$. The set of pure single-qubit stabilizer states, $\mathcal{S}(\mathcal{H}_2)$, are the ± 1 eigenstates of the Pauli operators;

$$\mathcal{S}(\mathcal{H}_2) = \{|\psi\rangle \,|\, P\,|\psi\rangle = \pm |\psi\rangle, P \in \mathcal{P}_1\} = \{|0\rangle, |1\rangle, |+\rangle, |-\rangle, |i+\rangle, |i-\rangle\},$$

where $|0\rangle, |1\rangle$ are the Z eigenstates, $|+\rangle, |-\rangle$ are the X eigenstates, and $|i+\rangle, |i-\rangle$ are the Y eigenstates. In this paper we use the density operator representation of these states.

35

Therefore we define $\Pi_{i,P} = |\psi\rangle\langle\psi|$, $\psi \in \mathcal{S}(\mathcal{H}_2)$ as the $i \in \{\pm 1\}$ eigenstate of $P \in \mathcal{P}_1$, for example $\Pi_{+1,Z} = |0\rangle\langle 0|$ and $\Pi_{-1,X} = |-\rangle\langle -|$.

Unitary transformations in the single-qubit stabilizer subtheory are given by the Clifford Group, which is defined as the normalizer of the Pauli group;

$$Cl_2 = \left\{ U \in SU(2) | \, UPU^\dagger = \pm P', \forall P \in \mathcal{P}_1 : P' \in \mathcal{P}_1 \right\}.$$

For example the Pauli operators are elements of the Clifford group, as $P_i P_j P_i = (-1)^{[i,j]} P_j$, as well as the Hadamard gate H and Phase gate P, which can be chosen as generators of the Clifford group, $Cl_2 = \langle X, Y, Z, H, P \rangle$.

Measurements in the single-qubit stabilizer formalism are given by the Pauli observables $\mathcal{P}_1$. Hence the projectors associated to each outcome of a Pauli observable P are the projectors $\{\Pi_{i,P}\}_i$.

2.3.4 The Single-Rebit Stabilizer Subtheory

The single-rebit stabilizer subtheory [25] is a strict subtheory of the single-qubit stabilizer subtheory, where we only include in the subtheory states with real representations in the computational basis. Therefore our set of states is;

$$\mathcal{R}(\mathcal{H}_2) = \{|0\rangle, |1\rangle, |+\rangle, |-\rangle\}.$$

The set of allowed transformations in the subtheory, $\mathcal{T}_R$, are the Clifford operations that map $\mathcal{R}(\mathcal{H}_2)$ to itself;

$$\mathcal{T}_R = \langle X, Y, Z, H \rangle.$$

Allowed measurements in the subtheory are the X and Z observables.

2.4 The Rebit Stabilizer Subtheory is Contextual

We now show the single-rebit subtheory admits only contextual ontological models. The proof proceeds in a similar manner to the proof of contextuality for the single-qubit stabilizer formalism [52]. First we define two operationally equivalent maps;

$$\Phi_1(\rho) = \frac{1}{2}(\rho + Y\rho Y), \tag{2.5}$$

$$\Phi_2(\rho) = \frac{1}{2}(X\rho X + Z\rho Z). \tag{2.6}$$

It should be noted that Φ_1 and Φ_2 are operationally equivalent *within* the rebit subtheory, but are not operationally equivalent for larger subtheories, for example they are not operationally equivalent in the single-qubit stabilizer subtheory. As far as we are aware, this is the first such proof to rely on operational equivalence defined only within a subtheory.

We then note, Assuming PNC implies the ontic state space of any model of the single-rebit stabilizer subtheory can be partitioned into 4 regions, similar to the 8 for the single-qubit stabilizer subtheory. We can associate each partition to the X and Z eigenstate it lies in the support of. Therefore each partition can be labeled by 2-bits denoted $(x, z) \equiv \Delta_x \cap \Delta_z$.

> **Theorem 2.4.1** *Every ontological model of the single-rebit stabilizer subtheory that is preparation non-contextual is transformation contextual, and every ontological model that is traditionally non-contextual is transformation contextual[a].*
>
> ―――――――――
>
> [a] Also known as Kochen-Specker non-contextuality.

Proof. Assuming PNC, the stochastic maps given by eqs. (2.5) and (2.6) act on each partition in the following way. Γ_{Φ_1} maps the partition (x, z) to (x, z) and $(x \oplus 1, z \oplus 1)$ with equal probability, and Γ_{Φ_2} maps the partition (x, z) to $(x, z \oplus 1)$ and $(x \oplus 1, z)$ with equal probability. Therefore Γ_{Φ_1} and Γ_{Φ_2} must be represented differently in any PNC ontological model. Implying preparation non-contextuality $\Rightarrow$ transformation contextuality.

Similarly the partitions can be defined by the observables in the subtheory when an assumption of outcome determinism (OD) is made. These partition are mapped as given above. Hence, as OD + MNC implies traditional non-contextuality [48] we have traditional non-contextuality $\Rightarrow$ transformation contextuality. Therefore, any ontological model of the singe-rebit stabilizer subtheory must be either preparation contextual or transformation contextual, and either transformation contextual or traditionally contextual. ∎

One interesting thing to note about this proof, is that for a rebit we do not need to rely on the structure of the allowed non-Pauli Clifford operations, i.e. only Pauli operations are required to demonstrate contextuality. For comparison, the single-qubit stabilizer subtheory requires a non-Pauli Clifford operation for any proof of contextuality, which can be inferred via the fact that the toy-model [77] is a universally non-contextual model[2].

―――――――――

[2] Which follows from its construction being equivalent to a particular choice of positively represented elements of the one-qubit Wigner function.

2.5 Frame Representations

Here we follow the usual quasi-probability representation (QPR) frame formalism [27].

Let $\mathcal{F} = \{F_\lambda | F_\lambda \in \mathbb{H}(\mathcal{H})\}$ and $\mathcal{D} = \{D_\lambda | D_\lambda \in \mathbb{H}(\mathcal{H})\}$, where $\mathbb{H}(\mathcal{H})$ is the set of Hermitian operators on $\mathcal{H}$, be a primal and dual frame such that;

$$A = \sum_\lambda \langle F_\lambda, A \rangle D_\lambda, \ \forall A \in \mathbb{H}(\mathcal{H}), \tag{2.7}$$

which we refer to as the *recombination formula*.

The condition for $\mathcal{F}$ and $\mathcal{D}$ to form a QPR is that they satisfy $\sum_\lambda F_\lambda = \mathbb{I}$ and $\mathrm{Tr}\,(D_\lambda) = 1$. Given a QPR we identify to every quantum state a representation in the primal frame $\mu_\rho(\lambda) = \langle F_\lambda, \rho \rangle$, and every POVM element a representation in the dual frame $\xi_{E_k}(\lambda) = \mathrm{Tr}\,(E_k D_\lambda)$[3]. We refer to ξ_{E_k} as measurement response functions, paralleling the definitions in the ontological models formalism. The inner product between the two representations, $\mu_\rho.\xi_{E_k}$, gives the quantum statistics;

$$\mathrm{Tr}(\rho E_k) = \mathrm{Tr}\left(\left(\sum_\lambda \langle F_\lambda, \rho \rangle D_\lambda\right) E_k\right),$$

$$= \sum_\lambda \langle F_\lambda, \rho \rangle \mathrm{Tr}\,(D_\lambda E_k) = \mu_\rho.\xi_{E_k},$$

where we have used eq. (2.7) to expand the density matrix ρ in the primal frame.

The representation of linear quantum maps, $\Phi \in \mathcal{L}(\mathcal{H})$, can be derived via;

$$\Phi(\rho) = \sum_\lambda \langle F_\lambda, \rho \rangle \Phi(D_\lambda),$$

$$\sum_\lambda \langle F_\lambda, \Phi(\rho) \rangle D_\lambda = \sum_{\lambda',\lambda} \langle F_\lambda, \rho \rangle \langle F_{\lambda'}, \Phi(D_\lambda) \rangle D_{\lambda'},$$

where in the first line we have applied eq. (2.7) to the LHS and linearity to the RHS, and have used eq. (2.7) to go from line 1 to 2. As this is a linear representation it must be the case that $\langle F_\lambda, \Phi(\rho) \rangle = \sum_{\lambda'} \langle F_\lambda, \Phi(D_{\lambda'}) \rangle \langle F_{\lambda'}, \rho \rangle$. Therefore we identify $\Gamma_\Phi(\lambda, \lambda') = \langle F_\lambda, \Phi(D_{\lambda'}) \rangle$ as the transition map representing Φ, in matrix notation we have $\mu_{\Phi(\rho)} = \Gamma_\Phi.\mu_\rho$. The transition map inherits normalization from the normalization of the primal frame $\sum_\lambda \langle F_\lambda, \Phi(D_{\lambda'}) \rangle = \langle \mathbb{I}, \Phi(D_{\lambda'}) \rangle = 1$, for CPTP Φ. Therefore we call Γ_Φ *quasi-stochastic*.

---

[3]The Hilbert-Schmidt inner-product is required in the definition of the primal or dual representations to reproduce the quantum statistics.

2.6 Any Quasi-Probability Representation of the Single-Qubit Stabilizer Subtheory must be Negatively represented

Now we prove that any QPR of the single-qubit stabilizer subtheory must be negatively represented. To do so we let, without loss of generality, an arbitrary QPR represent all states and all measurements positively. As is the case in the 4-state Wigner function [77, 79] and 8-state model [84, 11, 44]. Let the QPR be defined by a primal and dual frame $\mathcal{F}$ and $\mathcal{D}$. Therefore the QPR satisfies;

$$\left.\begin{array}{rl} \langle F_\lambda, \Pi_{i,P} \rangle & \geq 0 \\ \mathrm{Tr}\left(\Pi_{i,P} D_\lambda\right) & \geq 0 \end{array}\right\} \forall F_\lambda \in \mathcal{F}, \, \forall D_\lambda \in \mathcal{D}, \, \forall i \in \{\pm 1\}, \, \forall P \in \mathcal{P}_1. \tag{2.8}$$

From this we can prove several lemmas; Firstly, that the support of the response functions is constrained by the support of the state's representation;

> **Lemma 2.6.1** $\langle F_\lambda, \Pi_{i,P} \rangle > 0$ *implies* $\mathrm{Tr}\left(D_\lambda \Pi_{i,P}\right) = 1$ *and* $\mathrm{Tr}\left(D_\lambda \Pi_{-i,P}\right) = 0$.

Proof. By the Born rule we have;

$$\mathrm{Tr}\left(\Pi_{i,P} \Pi_{i,P}\right) = 1,$$
$$= \sum_\lambda \langle F_\lambda, \Pi_{i,P} \rangle \, \mathrm{Tr}\left(D_\lambda \Pi_{i,P}\right).$$

Therefore by equation 2.8 and normalization of the QPR, $\sum_\lambda \langle F_\lambda, \Pi_{i,P} \rangle = 1$, it must be the case that $\mathrm{Tr}\left(D_\lambda \Pi_{i,P}\right) = 1$, $\forall \lambda$ such that $\langle F_\lambda, \Pi_{i,P} \rangle > 0$.

Similarly we have by the Born rule;

$$\mathrm{Tr}\left(\Pi_{i,P} \Pi_{-i,P}\right) = 0,$$
$$= \sum_\lambda \langle F_\lambda, \Pi_{i,P} \rangle \, \mathrm{Tr}\left(D_\lambda \Pi_{-i,P}\right).$$

Therefore $\forall \lambda$ such that $\langle F_\lambda, \Pi_{i,P} \rangle > 0$ it must be the case that $\mathrm{Tr}\left(D_\lambda \Pi_{-i,P}\right) = 0$, again by equation 2.8. ∎

From this we can infer the supports of orthogonal eigenstates must be disjoint;

Corollary 2.6.2 $\langle F_\lambda, \Pi_{i,P} \rangle > 0$ *implies* $\langle F_\lambda, \Pi_{-i,P} \rangle = 0$.

Proof. For contradiction assume $\exists \lambda$ such that $\langle F_\lambda, \Pi_{i,P} \rangle > 0$ and $\langle F_\lambda, \Pi_{-i,P} \rangle > 0$. Lemma 2.6.1 implies $\text{Tr}\,(D_\lambda \Pi_{i,P}) = 1$, however by the born rule we have;

$$\text{Tr}\,(\Pi_{-i,P} \Pi_{i,P}) = 0,$$
$$= \sum_\lambda \langle F_\lambda, \Pi_{-i,P} \rangle \, \text{Tr}\,(D_\lambda \Pi_{i,P}) > 0.$$

Therefore we arrive at a contradiction. ∎

Secondly, it must be the case that every λ is in the support of $\Pi_{i,P}$ or $\Pi_{-i,P}$, for all P.

Lemma 2.6.3 *For all λ it must be the case that either $\langle F_\lambda, \Pi_{i,P} \rangle > 0$ or $\langle F_\lambda, \Pi_{-i,P} \rangle > 0$, $\forall P$.*

Proof. First we remove any F_λ from the primal frame that satisfies $\langle F_\lambda, \Pi_{i,P} \rangle = 0$, $\forall i, P$ because this implies $F_\lambda = 0$ by the definition of an inner product[4], and therefore is a trivial frame operator.

Therefore $\exists i, P$ such that $\langle F_\lambda, \Pi_{i,P} \rangle > 0$, $\forall \lambda$. From this we can infer that all other bases must have support on λ as we have;

$$\left\langle F_\lambda, \frac{\mathbb{I}}{2} \right\rangle = \frac{1}{2}(\langle F_\lambda, \Pi_{i,P} \rangle + \langle F_\lambda, \Pi_{-i,P} \rangle) > 0, \, \forall \lambda,$$
$$\implies \frac{1}{2}(\langle F_\lambda, \Pi_{i,P'} \rangle + \langle F_\lambda, \Pi_{-i,P'} \rangle) > 0, \, \forall P' \neq P.$$

Therefore by corollary 2.6.2 it must be the case that $\langle F_\lambda, \Pi_{i,P} \rangle > 0$ or $\langle F_\lambda, \Pi_{-i,P} \rangle > 0$ for all $P \in \mathcal{P}_1$. ∎

Lemma 2.6.3 also implies;

$$\langle F_\lambda, \Pi_{i,X} \rangle = \langle F_\lambda, \Pi_{j,Y} \rangle = \langle F_\lambda, \Pi_{k,Z} \rangle, \tag{2.9}$$

where i, j, k indicate which eigenbasis λ is in the support of.

---

[4]By noting that the set of projectors form a basis for the space.

Before proving the main theorem we note that as $D_\lambda \in \mathbb{H}(\mathcal{H})$ we have;

$$D_\lambda = \frac{1}{2}(\mathbb{I} + x_\lambda X + y_\lambda Y + z_\lambda Z), \tag{2.10}$$

$$= \frac{1}{2}\left(\mathbb{I} + x_\lambda(\Pi_{+1,X} - \Pi_{-1,X}) + y_\lambda(\Pi_{+1,Y} - \Pi_{-1,Y}) + z_\lambda(\Pi_{+1,Z} - \Pi_{-1,Z})\right), \tag{2.11}$$

where $x_\lambda, y_\lambda, z_\lambda \in \mathbb{R}$ and the $\frac{1}{2}$ on the identity is due to $\text{Tr}(D_\lambda) = 1$. However we can constrain the values $x_\lambda, y_\lambda, z_\lambda$ can take by considering lemma 2.6.1. Let $\lambda \in \text{supp}(\Pi_{i,P})$ therefore by lemma 2.6.1 we have;

$$\text{Tr}(D_\lambda \Pi_{i,P}) = 1,$$
$$= \text{Tr}\left(\frac{1}{2}(\mathbb{I} + x_\lambda X + y_\lambda Y + z_\lambda Z)\frac{\mathbb{I} + iP}{2}\right),$$
$$= \frac{1}{2}(1 + ip_\lambda) = 1,$$
$$\implies p_\lambda = i, \text{ as } i \in \{\pm 1\},$$

where p_λ is the phase on the Pauli operator P, i.e. if $P = X$ then $p_\lambda = x_\lambda$. Note this implies that the two 4-state Wigner functions and any QPR with the 8-state dual frame are the only possible QPR's of the single qubit stabilizer subtheory that represent all stabilizer states and all Pauli measurements positively. With this we can now prove the main theorem;

> **Theorem 2.6.4** *Any quasi-probability representation of the single-qubit stabilizer subtheory must be negatively represented.*

Proof. As previously we assume that all states and measurements are positively represented, equation 2.8. Further, without loss of generality we assume that the identity transformation is positively represented;

$$\Gamma_\mathbb{I}(\lambda', \lambda) = \langle F_{\lambda'}, D_\lambda \rangle \geq 0, \forall \lambda, \lambda'. \tag{2.12}$$

Substituting in equation 2.11 and applying lemma 2.6.3 and corollary 2.6.2, and using equation 2.9 we have;

$$\langle F_{\lambda'}, D_\lambda \rangle = \frac{1}{2}\left(\langle F_{\lambda'}, \Pi_{i_{\lambda'},X}\rangle + i_\lambda i_{\lambda'} \langle F_{\lambda'}, \Pi_{i_{\lambda'},X}\rangle + j_\lambda j_{\lambda'} \langle F_{\lambda'}, \Pi_{j_{\lambda'},Y}\rangle + k_\lambda k_{\lambda'} \langle F_{\lambda'}, \Pi_{k_{\lambda'},Z}\rangle\right),$$
$$= \frac{1}{2}(1 + i_\lambda i_{\lambda'} + j_\lambda j_{\lambda'} + k_\lambda k_{\lambda'}) \langle F_{\lambda'}, \Pi_{i_{\lambda'},X}\rangle \geq 0,$$
$$\implies (1 + i_\lambda i_{\lambda'} + j_\lambda j_{\lambda'} + k_\lambda k_{\lambda'}) \geq 0, \tag{2.13}$$

41

where we have used i_λ, $i_{\lambda'}$, etc. to indicate which eigenstate of each Pauli basis is in the support of λ and λ' respectively.

If we consider that each product $i_\lambda i_{\lambda'}$, $j_\lambda j_{\lambda'}$, and $k_\lambda k_{\lambda'}$ is ± 1 valued, we can see that at most two of the products can be negative, to enforce the positivity of eq. (2.13). Hence, the key insight of the proof is to notice that with access to the full Clifford group we can flip the sign of any product. Therefore for all λ and λ', $\exists C \in Cl_2$ such that $\langle F_{\lambda'}, CD_\lambda C^\dagger \rangle < 0$.

Finally all prior arguments in the paper apply equally if we let states be represented by the Hilbert-Schmidt inner product, $\mu_\rho(\lambda) = \text{Tr}(\text{F}_\lambda \rho)$, and PVMs by an arbitrary inner product, $\xi_\Pi(\lambda) = \langle D_\lambda, \Pi \rangle$. ∎

So for example, let $i_\lambda = i_{\lambda'} = j_\lambda = j_{\lambda'} = +1$, and $k_\lambda = k_{\lambda'} = -1$ which gives $i_\lambda i_{\lambda'} + j_\lambda j_{\lambda'} + k_\lambda k_{\lambda'} = 3$. Therefore by considering the action of the Hadamard gate, $HXH^\dagger = Z$, $HYH^\dagger = -Y$, and $HZH^\dagger = X$, and applying a similar procedure to equation 2.13, we have;

$$\Gamma_H(\lambda', \lambda) = \langle F_{\lambda'}, HD_\lambda H^\dagger \rangle,$$
$$= (1 + k_\lambda i_{\lambda'} - j_\lambda j_{\lambda'} + i_\lambda k_{\lambda'}) \langle F_{\lambda'}, \Pi_{i_{\lambda'},X} \rangle,$$
$$= -2 \langle F_{\lambda'}, \Pi_{i_{\lambda'},X} \rangle < 0.$$

Note, in the 4-state Wigner functions the condition that forces $\Gamma_{\mathbb{I}}$ to be positively represented is $i_\lambda j_\lambda = \pm k_\lambda$, which is satisfied by our example. Comparatively the 8-state model represents the identity transformation negatively, and so immediately fails to satisfy theorem 2.6.4.

2.7 The Rebit Stabilizer Subtheory must be Negatively Represented by any Quasi-Probability Representation

The proof proceeds slightly differently to that given in theorem 2.6.4 due to the structure of the single-rebit subtheory. Firstly we note that lemma 2.6.1, lemma 2.6.3 and corollary 2.6.2 apply equally to any QPR of the single-rebit subtheory. Secondly, rather than the frame and dual operators having to span $\mathbb{H}(\mathcal{H}_2)$, we only require they span the real, under a computational basis representation, plane of single-qubit Hermitian operators[5].

---

[5]For the full n-rebit subtheory construction see ref **??**.

Further, we can project any representation to this subspace, as any component outside this span has no effect on the QPR. From this, we can see that any dual operator D_λ can be expressed as;

$$D_\lambda = \frac{1}{2}(\mathbb{I} + x_\lambda X + z_\lambda Z), \tag{2.14}$$

$$= \frac{1}{2}\left(\mathbb{I} + x_\lambda(\Pi_{+1,X} - \Pi_{-1,X}) + z_\lambda(\Pi_{+1,Z} - \Pi_{-1,Z})\right), \tag{2.15}$$

where $x_\lambda, z_\lambda \in \mathbb{R}$ and the $\frac{1}{2}$ on the identity is due to $\mathrm{Tr}\,(D_\lambda) = 1$. However again we can constrain the values x_λ, z_λ can take by considering lemma 2.6.1. Let $\lambda \in \mathrm{supp}(\Pi_{i,P})$ therefore by lemma 2.6.1 we have;

$$\mathrm{Tr}\,(D_\lambda \Pi_{i,P}) = 1,$$

$$= \mathrm{Tr}\left(\frac{1}{2}(\mathbb{I} + x_\lambda X + z_\lambda Z)\frac{\mathbb{I} + iP}{2}\right),$$

$$= \frac{1}{2}(1 + ip_\lambda) = 1,$$

$$\implies p_\lambda = i, \text{ as } i \in \{\pm 1\},$$

where again p_λ denotes the phase on Pauli operator P, so if $P = X$ then $p_\lambda = x_\lambda$.

With this established we can now prove that any QPR the rebit subtheory which positively represents all states and measurements, must represent transformations negatively;

> **Theorem 2.7.1** *Any quasi-probability representation of the single-rebit stabilizer subtheory must be negatively represented.*

Proof. Again we assume a positive representation of the identity transformation,

$$\Gamma_\mathbb{I}(\lambda', \lambda) = \langle F_{\lambda',\lambda}, D_\lambda \rangle \geq 0.$$

As before, substituting in equation 2.11 and applying lemma 2.6.3 and corollary 2.6.2, and using equation 2.9 we have;

$$\langle F_{\lambda'}, D_\lambda \rangle = \frac{1}{2}\left(\langle F_{\lambda'}, \mathbb{I}\rangle + i_\lambda \langle F_{\lambda'}, X\rangle + j_\lambda \langle F_{\lambda'}, Z\rangle\right),$$

$$= \frac{1}{2}\left(\langle F_{\lambda'}, \Pi_{i_{\lambda'},X}\rangle + i_\lambda i_{\lambda'} \langle F_{\lambda'}, \Pi_{i_{\lambda'},X}\rangle + j_\lambda j_{\lambda'} \langle F_{\lambda'}, \Pi_{j_{\lambda'},Z}\rangle\right),$$

$$= \frac{1}{2}\left(1 + i_\lambda i_{\lambda'} + j_\lambda j_{\lambda'}\right)\langle F_{\lambda'}, \Pi_{i_{\lambda'},X}\rangle \geq 0,$$

where again we have used i_λ, $i_{\lambda'}$, j_λ, and $j_{\lambda'}$ to denote which eigenstate of X and Z, respectively, λ and λ' lie in the support of. Therefore by the assumption that all states are positively represented we have;

$$(1 + i_\lambda i_{\lambda'} + j_\lambda j_{\lambda'}) \geq 0. \tag{2.16}$$

From this we can see that at most one of the products $i_\lambda i_{\lambda'}$ or $j_\lambda j_{\lambda'}$ can be equal to -1. Further there must exist a λ and λ' such that one of the products satisfies $i_\lambda i_{\lambda'} = -1$ or $j_\lambda j_{\lambda'} = -1$, because if all products were positive this would imply that every λ was in the support of one of the $+1$ or -1 eigenstates of X and Z and therefore the QPR could not reproduce the quantum statistics.

Taking the λ and λ' such that one of the products is negative and given that the representation of a transformation is given by;

$$\Gamma_\Phi(\lambda', \lambda) = \langle F_{\lambda'}, \Phi(D_\lambda) \rangle.$$

We can flip the sign of any of the products in eq. (2.16) by applying the correct Pauli, say P, therefore $\Gamma_P(\lambda', \lambda) = \langle F_{\lambda'}, P D_\lambda P \rangle < 0$. ∎

For example, if for some given λ and λ', $i_\lambda i_{\lambda'} = 1$ and $j_\lambda j_{\lambda'} = -1$ then the representation of the Z operation would be negative;

$$\Gamma_Z(\lambda', \lambda) = \langle F_{\lambda'}, Z D_\lambda Z \rangle = \frac{1}{2} \left(\langle F_{\lambda'}, \mathbb{I} \rangle - i_\lambda \langle F_{\lambda'}, X \rangle + j_\lambda \langle F_{\lambda'}, Z \rangle \right),$$

$$\propto (1 - i_\lambda i_{\lambda'} + j_\lambda j_{\lambda'}) < 0.$$

We note that theorem 2.4.1 and theorem 2.7.1 apply equally to *any* rebit subtheory, i.e. the subtheories defined by considering real density matrices in a given Pauli basis. Despite each subtheory possibly having a different set of allowed non-Pauli Clifford operations. This is because the proofs we have given only require we have access to all Pauli operations, a feature shared by all rebit subtheories[6].

2.8   Discussion

In this paper we have shown that, like the single-qubit stabilizer subtheory, the rebit subtheories require a contextual ontological model when transformations are included in the

---

[6]We propose a nomenclature for these theories: ReX-bit, ReY-bit, and ReZ-bit.

operational theory. This is surprising due to these subtheories being strict subtheories of the single-qubit stabilizer subtheory and are comparable to the qudit stabilizer subtheory. Further we have shown that the single-qubit stabilizer subtheory and rebit subtheories must be represented by negative quasi-probability representations. Re-establishing the link between contextuality and negativity when transformations are included in the operational description of a subtheory.

The seeming ubiquity of contextuality once transformations are included in an operational description seriously damages the prospect of viewing contextuality as *the* resource for quantum computational speed-up. Especially considering the n-rebit stabilizer subtheory was specifically constructed [25] to be a non-contextual subtheory of the contextual n-qubit stabilizer subtheory. However, for the n-rebit stabilizer subtheory ($n > 1$) the contextuality can be confined to the transformations, unlike the n-qubit stabilizer subtheory ($n > 1$).

It remains an open question what the physical significance of transformation contextuality is, and how it relates to the more explored notions of preparation and measurement contextuality. The proofs given here and in Ref. [52] suggest that contextuality can be "moved" around within an ontological model. However there are no known transformation non-contextual ontological models of the subtheories investigated here and more generally. This poses a question, is such model possible or is transformation contextuality somehow more fundamental than preparation and measurement contextuality?

2.8.1   Acknowledgments

Chapter 3

Quasi-Probability Representations of Transformations

3.1   Chapter Preamble

In this relatively short chapter, which does not fit within the contexts of chapters 1 and 2, we shall investigate a possible research avenue given the results of chapters 1 and 2. Chapters 1 and 2 investigate transformation contextuality and negativity within subtheories of the single-qubit stabilizer subtheory. However, one may ask: are there more general results on transformation contextuality for larger subtheories and quantum theory? And how does this relate to negativity, given the previously established relationship between contextuality and negativity?

Here I will present a proof that any finite QPR cannot represent all unitaries positively. Comparatively it is relatively straight-forward to construct a QPR where either all states or all measurements are positively represented. The proof rests on an assumption that the identity channel is represented by the identity map in a QPR. While this assumption is physically plausible, it could be dropped with a more thorough investigation.

The work presented here is from my own research, but I have been in communication with Hammam Qassim regarding the details of the proofs and possible routes forward.

3.2 Introduction

The relationship between negativity and contextuality within the quasi-probability representation formalism and the ontological models formalism has long been established [78, 28]. With the requirement of a negative representation of quantum theory being equivalent to the requirement of a contextual ontological model of quantum theory. There has been some discussion with my colleagues whether this statement extends to subtheories of quantum theory, which I believe it does, but warrants further investigation.

The equivalence between negativity and contextuality has proven to be a powerful tool in the foundationalist tool box [78, 28, 84, 82, 39, 25, 71], and has lead to some remarkable insights. Especially when this equivalence is presented in the context of quantum computation, as it is in [82, 39, 25, 71]. However, the results presented in chapters 1 and 2 show us that there had been a missing component in research into negativity and contextuality. Namely, the role of transformations, which had often been relegated to asides. These results posed a more general question to me: How are transformations in full quantum theory represented? Can we represent all unitaries positively?

As we shall see in the following sections, we will see that this is not possible. All quasi-probability representations that represent the identity transformation by the identity matrix must represent *almost-all* unitaries negatively. The assumption that the identity transformation is represented by the identity matrix is physically well justified. However, following the results I present a possible route for removing this assumption, which would allow us to prove the main result for all quasi-probability representations.

3.3 Definitions

Here we follow the treatment of quasi-probability representations (QPR) given in Ref. [78]. Additionally, for the current proof we restrict ourselves to QPRs defined over a finite measurable space $|\Lambda| = d_r$, where the subscript is used to distinguish this dimension from the dimension of the Hilbert space, $|\mathcal{H}| = d$, the QPR is a representation of.

For every density operator $\rho \in D(\mathcal{H})$ we associate it to a real-valued function μ_ρ defined over Λ. Mathematically $\mu_\rho : \Lambda \to \mathbb{R}$ such that μ_ρ is normalized, $\sum_\lambda \mu_\rho(\lambda) = 1$, and therefore we refer to them as a quasi-probability distribution. Every CPTP map $\Phi \in \mathcal{L}(\mathcal{H})$ is associated to a quasi-stochastic map Γ_Φ, such that if $\Phi(\rho) = \rho'$ we have $\Gamma_\Phi \mu_\rho = \mu_{\rho'}$. The map is called quasi-stochastic if it satisfies $\sum_{\lambda'} \Gamma_\Phi(\lambda', \lambda) = 1$. A POVM, $\{E_k\}$, in a QPR is represented by a set of real-valued functions $\{\xi_{E_k}\}$, where each function is associated

to each element, E_k, of the POVM. All these functions satisfy a normalization constraint $\sum_k \xi_{E_k}(\lambda) = 1$.

For a QPR to reproduce quantum theory, it must be the case that;

$$\mathrm{Tr}\,(E_k \Phi(\rho)) = \sum_{\lambda',\lambda} \xi_{E_k}(\lambda')\Gamma_\Phi(\lambda',\lambda)\mu_\rho(\lambda). \tag{3.1}$$

Further we note, the set of possible QPRs is a strict subset of the set of super-operator representations, wherein normalization conditions have been imposed on the super-operator representation.

We say a QPR gives a *non-negative* representation of states, if for all $\rho \in D(\mathcal{H})$ we have $\mu_\rho \geq 0$. Similarly we say a QPR gives a non-negative representation of measurements, if for all POVMs $\{E_k\}$ and all k we have $\xi_{E_k} \geq 0$. As per the proof in Ref [78], a non-negative representation of states *and* measurements is impossible. However we can easily construct QPRs where either states or measurements are positively represented. In the following we will see that this is not the case for unitaries; all finite QPRs must represent almost all unitaries negatively.

Finally, we note that in a QPR, channel composition can be represented by the combined channel representation. So if we apply two transformations on a quantum system $\Phi_2(\Phi_1(\rho)) = \Phi_3(\rho)$[1] then we can represent this sequence of transformations in a QPR by the representation of the combined channel, i.e. $\Gamma_{\Phi_2}\Gamma_{\Phi_1} = \Gamma_{\Phi_3}$[2];

$$\Phi_2(\Phi_1(\rho)) = \rho' = \Phi_3(\rho),$$
$$\longrightarrow \sum_{\lambda',\lambda} \Gamma_{\Phi_2}(\lambda'',\lambda')\Gamma_{\Phi_1}(\lambda',\lambda)\mu_\rho(\lambda) = \mu_{\rho'}(\lambda''),$$
$$= \sum_\lambda \left(\sum_{\lambda'} \Gamma_{\Phi_2}(\lambda'',\lambda')\Gamma_{\Phi_1}(\lambda',\lambda) \right) \mu_\rho(\lambda),$$
$$= \sum_\lambda \Gamma_{\Phi_3}(\lambda'',\lambda)\mu_\rho(\lambda) = \mu_{\rho'}(\lambda'').$$

Therefore we have;

$$\Gamma_{\Phi_3}(\lambda'',\lambda) = \sum_{\lambda'} \Gamma_{\Phi_2}(\lambda'',\lambda')\Gamma_{\Phi_1}(\lambda',\lambda), \tag{3.2}$$

which is just the familiar matrix multiplication equation and so $\Gamma_{\Phi_3} = \Gamma_{\Phi_2}\Gamma_{\Phi_1}$.

---

[1]Note: the reverse direction is not always possible, not all quantum channels can be decomposed into a sequence of other channels. For example the amplitude damping channel.

[2]Assuming that the set of μ_ρ form a basis for the set of functions over Λ.

3.4 Finite Quasi-Probability Representations and the Representation of Unitary Transformations

Without loss of generality we assume the identity transformation is positively represented, i.e. $\Gamma_{\mathbb{I}} \geq 0$. Further we utilize the following proposition;

Proposition 3.4.1 *In any finite quasi-probability representation, which represents the identity transformation positively, the identity transformation is represented by the identity map;*

$$\Gamma_{\mathbb{I}} = \mathbb{I}_{d_r}$$

This proposition makes intuitive sense, i.e. the identity map can be thought of "do nothing map" and so should be represented by a trivial map. Further, if this is not the case then the QPR is tracking extraneous information that is not relevant for reproducing the desired statistics. In the next section we investigate proposition 3.4.1 and give an argument for imposing it. However, a conclusive proof of proposition 3.4.1 would be desirable or at least a weaker lemma that can still be used to prove the main theorem.

Armed with proposition 3.4.1 we can now prove the main theorem.

Theorem 3.4.2 *Any finite quasi-probability representation, which represents the identity map positively and satisfies proposition 3.4.1, cannot be non-negative for all transformations.*

Proof. Assume for contradiction that $\Gamma_U \geq 0$, $\forall U \in U(d)$. By the properties of a QPR we have;

$$\Gamma_U \Gamma_{U^\dagger} = \Gamma_{\mathbb{I}_d} = \mathbb{I}_{d_r}.$$

This implies that $\Gamma_{U^\dagger} = \Gamma_U^{-1}$.

By assumption both Γ_U and $\Gamma_{U^\dagger}$ are positively represented and therefore are represented by stochastic matrices. The only invertible stochastic matrices are permutation matrices this implies Γ_U and $\Gamma_{U^\dagger}$ are permutation matrices. Given that in all, there are $|\mathcal{S}_{d_r}| = d_r!$ permutation matrices, where $\mathcal{S}_{d_r}$ is the set of all $d_r \times d_r$ permutation matrices, it is clear that $|\mathcal{S}_{d_r}| < |U(d)|$ for all finite d_r and d. Therefore only finitely many unitaries can be positively represented, implying any QPR cannot be non-negative for transformations. ∎

Theorem 3.4.2 allows us to immediately prove two corollaries:

> **Corollary 3.4.3** *Any finite QPR, defined over $\mathbb{R}^{d_r}$, can only represent at most $d_r!$ unitaries positively.*

Proof. The size of the set of permutation matrices is $d_r!$. ∎

> **Corollary 3.4.4** *In any quasi-probability representation, defined over $\mathbb{R}^{d_r}$, **almost all** unitaries are negatively represented.*

Proof. As $|\mathcal{S}_{d_r}|$ is finite for all finite d_r and $|U(d)|$ is infinite for all finite d the set of positively represented unitaries is a measure-zero set of $U(d)$. ∎

3.5 Representing the Identity Transformation in QPRs

Proposition 3.4.1 can be physically justified as follows. If we perform no transformation on a physical system we expect its state to remain static, including its ontological states. Therefore the map should be represented by a "do nothing" map, in a QPR which aims to reproduce the dynamics of an ontological model. However, there may be the possibility of some kind of steady-state dynamics, which would be captured by the identity transformation. There is some ambiguity here though as steady-state dynamics are often consider to take place over some time scale, which means by subdividing the process many identity maps could be constructed. Hence this raises the question which map is "the" identity map, with the best contender again being the instantaneous 'do nothing' map.

However, the QPR formalism makes no assumption about the form of the identity map, but does imply some very stringent constraints on its representation. The first thing to note is that eq. (3.2) implies the identity transformation is idempotent and therefore a projector;

$$\Gamma_{\mathbb{I}}\Gamma_{\mathbb{I}} = \Gamma_{\mathbb{I}} \tag{3.3}$$

Secondly, all quasi-probability distributions in the QPR, μ_ρ, lie in the $+1$ eigenspace of the identity map as;

$$\Gamma_{\mathbb{I}}\mu_\rho = \mu_\rho, \ \forall \rho. \tag{3.4}$$

Thirdly, the identity transformation is the identity element for the set of transformations represented by the QPR, via eq. (3.2);

$$\Gamma_{\mathbb{I}}\Gamma_{\Phi} = \Gamma_{\Phi}\Gamma_{\mathbb{I}} = \Gamma_{\Phi}, \, \forall \Phi \tag{3.5}$$

Fourthly, all response functions ξ_{E_k} also lie in the $+1$ eigenspace of the identity map;

$$\xi_{E_k}\Gamma_{\mathbb{I}} = \xi_{E_k}, \, \forall E_k. \tag{3.6}$$

Lastly, as we are interested in positivity we can assume the identity transformation is positively represented;

$$\Gamma_{\mathbb{I}} \geq 0. \tag{3.7}$$

What do eqs. (3.3) to (3.7) tell us about the representation of the identity? That there is a subspace of the representation that is "unused" by the QPR, as all physically relevant quantities are contained in the $+1$ eigenspace of the identity transformation. Therefore once we project out this junk subspace we are left with a representation of the identity that is the identity matrix. However, this process necessarily redefines the QPR and therefore cannot be seen as constituting a proof of negativity for the larger QPR, yet.

There may be a way of extending theorem 3.4.2 to representations defined over a larger space, which can be mapped to the smaller representation. i.e, if we let a QPR, $\mathcal{Q}_1$, that satisfies proposition 3.4.1 be defined over $\mathbb{R}^{d_I}$ and consider another QPR, $\mathcal{Q}_2$, defined over $\mathbb{R}^{d_r}$, satisfying eqs. (3.3) to (3.7), such that the subspace defined by the projector $\Gamma_{\mathbb{I}}$ is of dimension $d_r - d_I$. Then can we show that the necessity of a negative representation of unitaries in $\mathcal{Q}_1$ implies a negative representation in $\mathcal{Q}_2$?

3.6   Discussion

By theorem 3.4.2 we can see that a QPR that satisfies proposition 3.4.1 cannot have a non-negative representation for all unitaries. This is in stark contrast to the case of states and measurements, which individually *do* admit non-negative QPRs. This raises the question, what is it about the structure of transformation in quantum theory that distinguish them from states and measurement?

Additionally, can the bound on the number of positively represented unitaries given by corollary 3.4.3 be tightened? There is some evidence to suggest that a tighter bound

does indeed exist. If we consider the single-qubit toy-model proposed by Spekkens [77], which satisfies proposition 3.4.1, only 12 unitaries are positively represented. Compared to the potential of $24 = 4!$ being positively represented, with the final 12 permutation matrices being assigned to anti-unitary operations. Similarly, the n-qudit Wigner function[3], defined over a space of size d^n, does not represent $d^n!$ unitaries positively.

Finally, there is some potential for additional insights over the previously know negativity $\Leftrightarrow$ contextuality results [78, 28], where previous results focused on the prepare-measure operational setting. If indeed it can be shown that no non-negative quasi-probability representation of unitaries can be achieved for the n-qubit Clifford group, then this would imply that negativity *is* necessary for universal quantum computation. Sadly the same can not be said of sufficiency due to the results presented in chapter 2.

---

[3]With odd-prime d.

PART III

The n-qubit Stabilizer Subtheory and Ontological Models Thereof

Chapter 4

Using the Wigner Function to Represent the n-Qubit Stabilizer Subtheory

4.1   Chapter Preamble

In this chapter I shall present a project that led into the papers covered in chapters 5 and 6. Joel J Wallman gave Hammam Qassim and I a project to investigate whether the n-qubit stabilizer formalism could be reproduced by a frame-switching Wigner function. This frame-switching model was intended to be an extension of the 8-state Wigner function and to parallel the n-qudit Wigner function, which gives an efficiently simulable classical weak simulation for qudits of odd-prime dimension.

The outcome of the project was that a frame-switching Wigner function was unable to represent stabilizer measurement update positively. This then lead to me stripping away the Wigner function machinery to eventually construct the ψ-epistemic model presented in chapter 5, and revealing the result presented in chapter 6. Hence, this chapter is designed to serve as in-depth introduction to the results in chapters 5 and 6, covering the background material more thoroughly. It is also my hope that this chapter will serve as a bridge between the recent result by R. Raussendorf et al. [69], a quasi-probability representation of the n-qubit stabilizer formalism, and the ψ-epistemic model of the n-qubit stabilizer formalism presented in chapter 5.

The results presented in this chapter are exclusively from my own research efforts, with

H. Qassim's efforts focusing on the simulation aspect of the project. However, we were both in regular contact about how our various directions were progressing, and so I consider this joint work with him.

4.2 Introduction

In this chapter, we will investigate the failure of a research project that led to the results presented in chapters 5 and 6. Namely, research into whether a frame-switching Wigner could be used as a basis for simulating the n-qubit stabilizer subtheory, and provide a construction for a ψ-epistemic model of the subtheory. As chapters 5 and 6 are reproductions of their respective papers, they are relatively light on explaining the background theory. Therefore, this chapter has been designed as a more detailed review of the necessary theoretical background, and compliment the material covered in the later chapters.

The n-qubit stabilizer subtheory forms the basis for many applications in quantum information research [31, 32]. However, at the time this project was started, there was no known classical model[1] of the subtheory, despite it admitting an efficient classical simulation. This was in stark contrast to the n-qudit stabilizer subtheory for odd prime d, which admits not only a classical phase-space representation but this representation is also non-contextual in all regards [82, 83, 39]. Therefore it was an open problem as to how this disconnect between the qubit and qudit stabilizer subtheories could be resolved. Especially considering that state-independent proofs of traditional contextuality could be constructed purely within the n-qubit stabilizer subtheory, for example Mermin-Peres square [55, 63, 56] and GHZ style proofs of contextuality [33]. Comparatively, no such proofs can be constructed for qudits of odd prime dimension, by the existence of a non-contextual ontological model.

With scalable quantum computation moving from the realm of theory to experiment, understanding the source of a quantum computer's power compared a classical counterpart has become even more relevant. It had been the hope within the foundations community that this quantum speed-up could be attributed to previously defined notions of non-classicality, such as contextuality and non-locality [7, 43, 55, 63, 56, 75]. Indeed, there are significant results showing the necessity, and sometimes sufficiency, of these principles for a variety quantum information tasks [4, 70, 82, 68, 83, 9, 61, 3, 21].

However, no known definition of non-classicality has been able to identify *the* resource behind quantum computation, with contextuality arguably the closest. Hence, one of aims

---

[1]Where by "classical" models I essentially mean models that are not ψ-ontic.

of this project was to categorize the contextuality within the n-qubit stabilizer formalism, to allow us to refine the definition of contextuality such that it could be identified as the crucial component of a quantum computer. Ironically, the work stemming from the project presented in this chapter has actually suggested that another, more contentious, definition of non-classicality, ψ-onticity, may be the resource behind quantum computation[2].

Recently R. Raussendorf et al. [69] presented a quasi-probability representation that *can* represent the n-qubit stabilizer formalism. Their approach is very similar to the frame-switching Wigner function work presented here, with both being based on quasi-probability representations with overcomplete frames. However, their construction, to avoid the results of chapter 6, is remarkably similar to the ψ-epistemic model presented in this thesis. It is my conjecture that the ψ-epistemic model and overcomplete-frame representation from Raussendorf et al. are actually equivalent up to some relaxation[3]. Hence, this chapter is intended as both an introduction to the work in chapters 5 and 6, and as a reference point for any work that may link the ψ-epistemic model in chapter 5 to the over-complete frame representation presented by Raussendorf et al. I have presented my thoughts on how this connection manifests in the final section of this chapter.

The chapter is structured as follows; To begin, I provide an in-depth review of the theoretical machinery of n-qubit stabilizer formalism, covering all operationally relevant elements of the theory: states, transformations, measurement outcomes, and measurement updates. However, in the interest of space and your time I have omitted discussing the error-correction and quantum information tasks based on the formalism, but for the interested reader I suggest Daniel Gottesman's in depth book on the subject [32]. I then present the binary-symplectic representation of the n-qubit stabilizer formalism as it ubiquitously used within the literature, and provides a computational method to represent the stabilizer formalism[4]. From this, I briefly cover the celebrated Gottesman-Knill theorem, as it provides much of the motivation behind the project. From this we will move on to quasi-probability representations and frames, of which the Wigner function is an example. A detailed review of the frame formalism of quasi-probability representations is given, which is used extensively by the frame-switching Wigner function. This then naturally leads us to defining the original discrete Wigner function.

With the background material introduced we will then move onto the results of the project. Firstly, I will define the frame-switching Wigner function, and motivate why we might be interested in such a construction. We will also verify that the construction forms a

---

[2]See chapter 6 for more details.

[3]Once the model's ontology is reduced as detailed in the appendix of chapter 6 and the over-complete frame is restricted to a particular set of phase-point operators.

[4]Which is used in the 2-qubit mathematica code accompanying chapter 5, attached in the appendix

valid quasi-probability representation. Secondly, I provide a set of positivity theorems and one crucial negativity theorem. Theorem 4.4.3 gives us an *iff* condition defining which stabilizer states are positively represented by a given Wigner function. Theorem 4.4.5 proves that Clifford transformations have positively represented transition matrices, if they are composed with a *frame switching* operation. Theorem 4.4.6 then tells us that all stabilizer measurement projectors are positively represented in *any* frame. Lastly, theorem 4.4.8 proves that for any given frame there always exists a negatively represented measurement update map, for some choice of stabilizer measurement. Finally, I derive the representation of the β-function in the binary-symplectic representation, and provide a qualitative link between my work and the work of Raussendorf et al.

4.3 Preliminaries

4.3.1 The n-Qubit Stabilizer Subtheory

The n-qubit stabilizer formalism can be constructed from the n-qubit Pauli group;

Definition 4.3.1 *The n-qubit* **Pauli group**, $\mathcal{P}_n$, *is the set of Pauli operators;*

$$\mathcal{P}_n := \left\{ i^p \bigotimes_{i=1}^{n} P_i \middle| P_i \in \{\mathbb{I}, X, Y, Z\}, p \in \mathbb{Z}_4 \right\}$$

Where $\mathbb{I}$ is the identity operator and X, Y and Z are the single-qubit Pauli operators.

The above single-qubit Pauli operators can be expressed in the computational basis as;

$$\mathbb{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

For the purposes of this chapter it will prove easier to define the n-qubit stabilizer formalism via the *projective Pauli group*;

Definition 4.3.2 *The n-qubit **projective Pauli group**, $\widetilde{\mathcal{P}}_n$, is constructed via modding out the global phases in the n-qubit Pauli group, $\mathcal{P}_n$;*

$$\widetilde{\mathcal{P}}_n = \{\otimes_{i=1}^n P_i \,|\, P_i \in \{\mathbb{I}, X, Y, Z\}\} = \mathcal{P}_n \backslash U(1) \tag{4.1}$$

Stabilizer States

To define a stabilizer state we first need to define the notion of a *stabilizer operator*. A stabilizer operator, S, is an operator that leaves a state, $|\psi\rangle$, invariant when it is applied to a system, i.e., $S|\psi\rangle = |\psi\rangle$ and therefore $|\psi\rangle$ is in the $+1$ eigenspace of S. The motivation for working with the projective Pauli group rather than the full Pauli group is that the full Pauli group contains operators that stabilize no state[5]. Therefore we select the set of operators from the Pauli group which stabilize quantum states. This turns out to be all elements of the projective Pauli group with a ± 1 phase, where we cut $-\mathbb{I}$ as it stabilizes no state.

Definition 4.3.3 *The set of n-qubit **stabilizer operators**, $\mathcal{S}_n$, is all projective Pauli elements with a ± 1 phase, minus $-\mathbb{I}$;*

$$\mathcal{S}_n = \left\{\pm P \,|\, P \in \widetilde{\mathcal{P}}_n\right\} \backslash -\mathbb{I} \cong (\mathbb{Z}_2 \times \widetilde{\mathcal{P}}_n) \backslash -\mathbb{I}.$$

To define the set of pure stabilizer states we note that all the projective Pauli operators are rank-2 and therefore have eigenspaces with dimension 2^{n-1}. Therefore a single pure stabilizer state can be in the $+1$ eigenspaces of many stabilizer operators, and furthermore the set of stabilizer operators of which it is a $+1$ eigenvector uniquely identifies a pure stabilizer state.

Let us denote the set of stabilizer operators that stabilizes some state $|\psi\rangle$ by;

$$\mathcal{S}(\psi) = \{S \,|\, S|\psi\rangle = |\psi\rangle \,;\, S \in \mathcal{S}_n\}. \tag{4.2}$$

Firstly we can see that all elements of $\mathcal{S}(\psi)$ must mutually commute as follows. Given any

---

[5]For example, the Pauli operator iX satisfies $(iX)^2 = -\mathbb{I}$ and therefore stabilizes no state. To see this assume iX stabilizes an arbitrary state $|\psi\rangle$, we then have $|\psi\rangle = iX|\psi\rangle = iXiX|\psi\rangle = -|\psi\rangle$ and thus reach a contradiction.

$S_1, S_2 \in \mathcal{S}(\psi)$ we have:

$$|\psi\rangle = S_1 |\psi\rangle = S_1 S_2 |\psi\rangle ,$$
$$|\psi\rangle = S_2 |\psi\rangle = S_2 S_1 |\psi\rangle ,$$

and ss Pauli operators either commute or anti-commute, S_1 and S_2 must commute. Secondly, it is straight forward to see that if $S_1 \in \mathcal{S}(\psi)$ and $S_2 \in \mathcal{S}(\psi)$ then $(S_1 S_2)$ must also stabilize $|\psi\rangle$ and therefore $(S_1 S_2) \in \mathcal{S}(\psi)$. Thirdly, as $\mathbb{I}$ stabilizes all states, $\mathbb{I} \in \mathcal{S}(\psi)$, $\forall \psi \in \mathcal{H}$. Finally all stabilizer operators satisfy $S^2 = \mathbb{I}$ and therefore S are self-adjoint operators, i.e., $S = S^{-1}$ since they are also unitary. Putting all these inferences together we can see that $\mathcal{S}(\psi)$ is an *Abelian group*[6]. Furthermore the stabilizer groups of pure stabilizer states are the largest Abelian subgroups of the Pauli group and are specified by n generators [32], i.e., $|\mathcal{S}(\psi)| = 2^n$. We refer to the stabilizer groups of pure stabilizer states as *maximal* Abelian subgroups of the Pauli group, of which there are [2];

$$|\{\mathcal{S}(\mathcal{H}_{2^n})\}| = 2^n \prod_{k=0}^{n-1}(2^{n-k} + 1) = 2^{(.5+o(1))n^2} .$$

Throughout this part, chapter, and indeed thesis, I will employ a variety of methods to represent stabilizer groups, and more generally sets of Pauli operators. Here we will introduce a few key methods. The first, and most straight-forward, is to simply treat the stabilizer groups as a set, i.e.;

$$\mathcal{S}(\psi) = \{S | S |\psi\rangle = |\psi\rangle , S \in \mathcal{S}_n \} . \tag{4.3}$$

The second, more explicit, method is to express the group elements in terms of elements of the projective Pauli group, which is used extensively throughout chapters 5 and 6;

$$\mathcal{S}(\psi) = \left\{ (-1)^{\gamma(b)} P_b | P_b \in \widetilde{\mathcal{P}}_n \right\} , \tag{4.4}$$

where I have omitted the usual conditions for simplicity. This representation is far more useful when we are looking at how stabilizer groups update after a measurement, which will be covered below. The final method is to refer directly to a choice of generators of the group:

$$\mathcal{S}(\psi) = \langle G_1, G_2, ..., G_n \rangle . \tag{4.5}$$

---

[6]Where associativity is inherited from the fact that stabilizer operators are linear operators on $\mathcal{H}$.

Again this representation is useful for understanding how stabilizer groups update through measurement.

We can also reverse the direction of representation and use stabilizer groups and generators to represent the density matrices of stabilizer states;

$$\rho_\psi = \prod_{i=1}^{n} \frac{\mathbb{I} + G_i}{2} = \frac{1}{2^n} \sum_{S \in \mathcal{S}(\psi)} S, \tag{4.6}$$

where $\frac{\mathbb{I}+G_i}{2}$ projects on to the $+1$ eigenspace of G_i. The stabilizer groups can also give us information on the inner-product between two stabilizer states, the most important for us is the condition for two stabilizer states to be orthogonal;

Theorem 4.3.1 *Two pure stabilizer states, ψ and ϕ, are orthogonal, i.e. $\langle \psi | \phi \rangle = 0$, iff $\exists P \in S(\psi)$ such that $-P \in S(\phi)$.*

Proof. For the first direction assume $\exists P \in S(\psi)$ such that $-P \in S(\phi)$: Therefore,

$$\langle \psi | \phi \rangle = \langle \psi | P | \phi \rangle = - \langle \psi | \phi \rangle,$$
$$\Rightarrow \langle \psi | \phi \rangle = 0.$$

For the second direction assume $|\langle \psi | \phi \rangle|^2 = 0$: Define the ψ and ϕ's stabilizer groups as;

$$S(\psi) = \{ (-1)^{p_s} P_s | (-1)^{p_s} P_s | \psi \rangle = | \psi \rangle \,;\, P_s \in \widetilde{\mathcal{P}}_n, p_s \in \{0,1\} \},$$
$$S(\phi) = \{ (-1)^{p_t} P_t | (-1)^{p_t} P_t | \phi \rangle = | \phi \rangle \,;\, P_t \in \widetilde{\mathcal{P}}_n, p_t \in \{0,1\} \}.$$

We can express the inner product as;

$$|\langle \psi | \phi \rangle|^2 = \mathrm{Tr}\left(\rho_\psi \rho_\phi \right),$$
$$= \frac{1}{d^2} \sum_{s \in S(\psi), t \in S(\phi)} \mathrm{Tr}\left((-1)^{p_s + p_t} P_s P_t \right),$$

where we have used the stabilizer state representation $\rho_\psi = \frac{1}{d} \sum_{s \in S(\psi)} (-1)^{p_s} P_s$ and $\rho_\phi = \frac{1}{d} \sum_{t \in S(\phi)} (-1)^{p_t} P_t$. For Projective Pauli operators we apply the identity $\mathrm{Tr}\left(P_s P_t \right) = d\delta_{s,t}$;

$$|\langle \psi | \phi \rangle|^2 = \frac{1}{d} \sum_{s=t} (-1)^{p_s + p_t}. \tag{4.7}$$

60

So for the sum in eq. (4.7) to equal zero it must be the case that the vector $v = (..., p_s + p_t, ...)_{s=t}$ is balanced (i.e. it has equal 0's and 1's). For this to be the case there must exists a $s = t$ such that $p_s \neq p_t$, i.e. $p_s = p_t + 1 \bmod 2$. Therefore, $\langle \psi | \phi \rangle = 0 \Rightarrow \exists P \in \mathcal{S}(\psi)$ and $-P \in \mathcal{S}(\phi)$. ∎

Further, eq. (4.7) also allows us to work out the overlap between two stabilizer states;

Corollary 4.3.2 *Two non-orthogonal pure stabilizer states $|\psi\rangle$ and $|\phi\rangle$ have overlap;*

$$|\langle \psi | \phi \rangle|^2 = 2^{-(n-s)}, \tag{4.8}$$

where 2^s is the size of the largest mutual subgroup of $\mathcal{S}(\psi)$ and $\mathcal{S}(\phi)$, i.e. $M = \mathcal{S}(\psi) \cap \mathcal{S}(\phi)$ and $|M| = 2^s$.

Proof. As $|\psi\rangle$ and $|\phi\rangle$ are non-orthogonal by assumption this implies that $p_s = p_t$, $\forall s, t \in \mathcal{S}(\psi) \cap \mathcal{S}(\phi)$ in eq. (4.7). Therefore the sum counts the number of elements shared between $\mathcal{S}(\psi)$ and $\mathcal{S}(\phi)$, which is given by the largest mutual subgroup, which we define to have size 2^s, i.e. this subgroup is specified by s generators. ∎

Lastly, for mixed stabilizer states there are two classes of mixed states we might consider. The first are simple probabilistic mixtures of pure stabilizer states;

$$\mathcal{S}(D(\mathcal{H}_{2^n})) = \left\{ \rho \,|\, \rho = \sum_{\{|\psi_i\rangle\}} p_i \rho_{\psi_i}; \, 0 \leq p_i \leq 1, \sum_i p_i = 1, \exists \{|\psi_i\rangle\} \subset \mathcal{S}(\mathcal{H}_{2^n}) \right\},$$

where $\mathcal{S}(\mathcal{H}_{2^n})$ is the set of pure stabilizer states, and $D(\mathcal{H}_{2^n})$ the set of density operators associated with $\mathcal{H}_{2^n}$. It is clear the set of states $\mathcal{S}(D(\mathcal{H}_{2^n}))$ is infinite for any n. The second set of mixed stabilizer states are what I call *sub-normalized* stabilizer states, that is mixed states that admit a stabilizer group;

$$\mathcal{SN}(D(\mathcal{H}_{2^n})) = \left\{ \rho \,|\, \rho = \prod_{i=1}^{k} \frac{\mathbb{I} + G_i}{2}; \, k < n \right\},$$

where $\{G_i\}$ is a set of independent generators. As $\mathcal{SN}(D(\mathcal{H}_{2^n})$ is finite, due to $\mathcal{S}_n$ being finite, we know it is therefore a measure-zero set of $\mathcal{S}(D(\mathcal{H}_{2^n}))$. It is in fact this set of density operators that forms the basis for all error-correcting codes, as these operators represent the code-space of the code.

61

Clifford Operations

For stabilizer circuits, composed of input stabilizer states and transformations on them, we require that any transformation in the circuit only maps stabilizer states to other stabilizer states. Similarly with measurements we require that the post-measurement state is a stabilizer state. The set of such transformations is the Clifford group and the possible measurements are the Pauli observables.

Considering we can express a stabilizer state in terms of its stabilizer group, we want transformations in stabilizer circuits to map stabilizer operators to stabilizer operators. In group theory the set of such operations is known as the normalizer of the group. This allows us to define the stabilizer operations as follows;

> **Definition 4.3.4** *The set of n-qubit stabilizer transformations is the Clifford group, which is defined as the normalizer of the Pauli group;*
>
> $$Cl_{2^n} := \left\{ U \,|\, UPU^\dagger = P', \forall P \in \mathcal{S}_n; \, P' \in \mathcal{S}_n, \, U \in U(\mathcal{H}_{2^n}) \right\},$$
>
> *where $U(\mathcal{H}_{2^n})$ is the set of unitary operations on $\mathcal{H}_{2^n}$, and $\mathcal{S}_n$ is the set of stabilizer operators.*

Further, any stabilizer state can be constructed from the computational all-zero state, $|0\rangle^{\otimes n}$, by the application of an appropriate Clifford, i.e.

$$\forall |\psi\rangle \in \mathcal{S}(\mathcal{H}_{2^n}), \, \exists C \in CL_{2^n} \text{ such that } C\,|0\rangle^{\otimes n} = |\psi\rangle \,.$$

Indeed for a given stabilizer state there are *many* Clifford operations that map the all-zero state to a given stabilizer state.

In most error correction schemes the set of Clifford operations that are used to generate the full Clifford group are the single qubit Hadamard gate, H_i, the single qubit Phase gate, S_i, and either the two-qubit CNOT gate, $CNOT_{ij}$, or the controlled-Z gate, CZ_{ij}, i.e.

$$Cl_{2^n} = \langle H_i, S_i, CNOT_{ij} \rangle_{i,j} = \langle H_i, S_i, CZ_{i,j} \rangle_{i,j}, \tag{4.9}$$

where the indices i and j indicate which qubit the operation is acting on. For the CNOT and CZ gates i is the control qubit and j the target qubit. In the computational basis

these operations have matrix representations given by;

$$H = \tfrac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \qquad S = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix},$$

$$CNOT = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad CZ = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

It should be noted that all single qubit Paulis can be generated by the Hadamard and Phase gates, as $S^2 = Z$, $HZH = X$, and $SXS^\dagger = Y$. While the CNOT gate is the canonical standard entangling gate used in the stabilizer formalism, the controlled-Z gate has the nice property of being symmetric on the target and control qubits, and therefore is slightly easier to work with. Interestingly, the addition of *any* non-Clifford gate to the Clifford group promotes the Clifford group to quantum computational universality [57, 58].

In the context of this project it is useful to note that this definition implies that the Clifford group is uniquely determined by its action on the projective Pauli group. Further, a Clifford operation's action on the Pauli group can be described in terms of its action on a generating set of the Pauli group, and as the Clifford group represents an automorphism on the Pauli group it maps generating sets to generating sets. For use later, let us define the notation for the action of a Clifford operation on a generic projective Pauli operator as follows;

Definition 4.3.5 *The action of a Clifford operation, C, on a projective Pauli operator, $P_i \in \widetilde{\mathcal{P}}_n$, is given by;*

$$CP_iC^\dagger = (-1)^{\phi_C(j)}P_j$$

Where $\phi_C(j) \in \mathbb{Z}_2$, $j = C(i)$ and $P_j \in \widetilde{\mathcal{P}}_n$.

The reason why we index the phase by j rather than i is to simplify the mathematics when dealing with phase-point operators.

Stabilizer Measurements

The allowed measurements in stabilizer circuits are the stabilizer observables, which can be considered to be the Projective Pauli group elements with some additional classical post-processing. For this reason, I focus mainly on representing measurements of the projective Pauli observables;

> **Definition 4.3.6** *The set of allowed measurements in an n-qubit stabilizer circuit are the Projective Pauli observables $\widetilde{\mathcal{P}}_n$ with classical post-processing bit flips;*
>
> $$M_{STAB} \cong \mathbb{Z}_2 \times \widetilde{\mathcal{P}}_n. \tag{4.10}$$

As the Projective Pauli observables are binary outcome measurements, with ± 1 eigenvalues, we can write the projector onto the $(-1)^k$, $k \in \{0, 1\}$ eigenspace of a projective Pauli observable, $P \in \widetilde{\mathcal{P}}_n$, as;

$$\Pi_{k,P} = \frac{\mathbb{I} + (-1)^k P}{2}.$$

This expression is derived from the identities $P = \Pi_{+1,P} - \Pi_{-1,P}$ and $\mathbb{I} = \Pi_{+1,P} + \Pi_{-1,P}$.

The operational statistics of the stabilizer formalism can easily be expressed by looking at the expectation values of a projective Pauli observable, given an input stabilizer state;

> **Lemma 4.3.3** *Given an input stabilizer state $|\psi\rangle \in \mathcal{S}(\mathcal{H}_{2^n})$, with stabilizer group $\mathcal{S}(\psi)$, and projective Pauli observable $M \in \widetilde{\mathcal{P}}_n$, the expectation value of the observable is given by;*
>
> $$\langle M \rangle_\psi = \begin{cases} 1: & M \in \mathcal{S}(\psi) & (C1), \\ 0: & \pm M \notin \mathcal{S}(\psi) & (C2), \\ -1: & -M \in \mathcal{S}(\psi) & (C3). \end{cases} \tag{4.11}$$

Proof. We have;

$$\langle M \rangle_\psi = \text{Tr}(M\rho_\psi),$$

$$= \frac{1}{2^n} \sum_{P \in \widetilde{\mathcal{S}}(\psi)} (-1)^p \text{Tr}(MP),$$

$$= \sum_{P \in \widetilde{\mathcal{S}}(\psi)} (-1)^p \delta_{M,P},$$

where we have used the identity $\text{Tr}(P_i P_j) = 2^n \delta_{i,j}$, $\forall P_i, P_j \in \widetilde{\mathcal{P}}_n$, and lower case p indicates the phase on the stabilizer element P. Therefore if $\pm M \in \mathcal{S}(\psi)$ we have $\langle M \rangle_\psi = \pm 1$, and 0 otherwise. ∎

Given that the stabilizer measurements are binary outcome and have eigenvalues of ± 1 we simply state the outcome statistics of a stabilizer circuit by referring to stabilizer groups: If M is in the pre-measurement stabilizer group then the outcome is given by the phase on the corresponding stabilizer element, and otherwise the outcome is uniformly random.

Stabilizer Measurement Update

Stabilizer measurement update rules can be derived via considering how a stabilizer state updates after the measurement of a Pauli observable. We can directly derive how a stabilizer group updates under measurement using the von Neumann-Lüder's update rule. If we measure some pure stabilizer state $\rho_\psi = |\psi\rangle \langle\psi|$ with measurement $M \in \widetilde{\mathcal{P}}_n$ and observe outcome $k \in \{0, 1\}$ then the post-measurement density matrix $\rho_{\psi'}$ is given by, with appropriate renormalization;

$$\rho_{\psi'} \propto \frac{\mathbb{I} + (-1)^k M}{2} \rho_\psi \frac{\mathbb{I} + (-1)^k M}{2},$$

$$\propto \sum_{S \in S(\psi)} S + (-1)^k (MS + SM) + MSM,$$

$$\propto \sum_{S \in S(\psi)\,|\,[S,M]=0} S + (-1)^k MS. \tag{4.12}$$

One of the very nice features of the Pauli group is the fact that given any Abelian subgroup of the Pauli group, $G \subset \mathcal{P}_n$, any Pauli operator commutes with exactly half of the elements

of G, and this subset forms another Abelian subgroup of $\mathcal{P}_n$, *or* every element of G. Therefore the final line in eq. (4.12) can be thought of as follows;

If M was an element of the pre-measurement stabilizer group then stabilizer state will stay the same, as will it's stabilizer group. To see this, consider that the sum is over all elements of the group and that both terms in the sum generate all elements of the group. So the sum can be collapsed to a sum over the group elements[7]. If M was not an element of the pre-measurement group then the sum will be over half the elements of the pre-measurement group that did commute with M, including the identity. Therefore we can see that the first term S is a sum over elements common to the pre and post measurement groups and the second term $(-1)^k MS$ identifies the new elements of the post-measurement stabilizer group. I.e. We remove elements of the group that do not commute with M, then add to the group M and compositions of M with the elements that do commute with M.

Looking at the projective representation of the pre-measurement stabilizer group we have;

$$S(\psi) = \left\{ (-1)^{\gamma(b)} P_b \right\}_b \xrightarrow{M,k}$$
$$S(\psi') = \left\{ (-1)^{\gamma(b)} P_b, (-1)^{\gamma(b)+k+\beta(b,m)} P_{b+m} \right\}_{b|\,[b,m]=0}, \qquad (4.13)$$

where we have used the definition $P_a P_b = (-1)^{\beta(a,b)} P_{a+b} |\,[a.b] = 0$, and an arbitrary index over elements of the Pauli group. One such index, the Binary-symplectic representation, will be explored in the next section.

There is one last way we can express the stabilizer measurement update rules, by referring to generators of the stabilizer groups. Suppose we have a pre-measurement stabilizer group given by;

$$S(\psi) = \langle G_1, G_2, \ldots, G_n \rangle. \qquad (4.14)$$

If we measure M with outcome k, we can update this set of generators to a new set of generators that generate the post-measurement stabilizer group in the following way; Check through all G_i to find the first G_i that doesn't commute with M, remove this G_i from the set of generators and then add $(-1)^k M$ in it's place, letting this $G_i = H$. Continue to find the set of G_i that do not commute with M. For every instance of G_i that does not commute with M, replace G_i with $G_i H$, which is guaranteed to commute with M as $[G_i H, M] = G_i H M - M G_i H = G_i H M - G_i H M = 0$. Furthermore, these new generators will also be independent of any other generator left in the generating set.

---

[7]Noting that by lemma 4.3.3 the outcome of M is equal to it's phase if it is in the group.

We can further abstract the above update rules. Given any stabilizer group $S(\psi)$ and any Projective Pauli operator $M \in \widetilde{\mathcal{P}}_n$ we can always express the group as;

$$S(\psi) = \langle G_1, G_2, \ldots, G_{n-1}, h \rangle, \tag{4.15}$$

such that $[G_i, M] = 0, \forall G_i$ and $\pm M \notin \langle G_1, G_2, \ldots, G_{n-1} \rangle$, and either we have $h = (-1)^k M$ or $[h, M] \neq 0$. Under this representation, updating the generators after a measurement of M, with outcome k, is simply be performed via removing h and adding $(-1)^k M$ in it's place;

$$S(\psi') = \langle G_1, G_2, \ldots, G_{n-1}, (-1)^k M \rangle. \tag{4.16}$$

This way of thinking about stabilizer groups and updating stabilizer groups after measurement is relevant for the work presented in both chapters 5 and 6. However is less relevant for the material presented in this chapter.

4.3.2   The Binary-symplectic Representation

Throughout the this chapter and the next two I will often refer to arbitrary indices over elements of the Projective Pauli group. This abstraction, while mathematically useful, does not lend itself well to computational investigations. For the purposes of such computations the *binary-symplectic representation* provides an expression of the stabilizer formalism that is easily implementable in computer software. In this section I will give a brief overview of the representation and in a following section derive an expression for the β-function in the binary-symplectic representation.

The binary-symplectic representation uses the group properties of the stabilizer theory to reduce the representation of stabilizer elements from quantum operators to binary vectors. This allows certain properties of the stabilizer formalism to be calculated without referring to the operator framework[8].

Any n-qubit stabilizer operator can be represented as a $2n + 1$ binary vector, where the $2n$ bits represents which members of a generating set for the Pauli group generates the given stabilizer operator, and the final bit determines the ± 1 phase of the stabilizer operator. So suppose we have some generating set, $\mathcal{G} = \{G_i\}_{i=1,\ldots,2n}$, for the Pauli group;

$$\langle G_i \rangle_{i=1,\ldots,2n} = \mathcal{P}_n.$$

---

[8]One major exception to this is calculating the statistics of circuits which contain *non*-stabilizer elements.

Then any stabilizer operator can be expressed as;

$$S = (-1)^s i^{f(\{g_i\})} \prod_{i=1}^{2n} G_i^{g_i}, \ \forall S \in \mathcal{S}_n \,|\, s, g_i \in \mathbb{Z}_2, \tag{4.17}$$

where we note that the product is necessarily ordered as the G_i do not mutually commute. From this we can represent any stabilizer operator as a binary vector;

$$\mathcal{B}(S) = (g_1, g_2, ..., g_{2n}|s).$$

In standard quantum information literature it is common practice to take the generating set to be the single-qubit X and Z operators, so;

$$\mathcal{G}_{BS} = \{X_i\}_{i=1,..,n} \cup \{Z_i\}_{i=1,...,n}, \tag{4.18}$$

where $X_i = \mathbb{I}_{j\neq i} \otimes X$ and $Z_i = \mathbb{I}_{j\neq i} \otimes Z$ are used to suppress tensor notation. With this generating set the binary vector representing a stabilizer operator can be broken into an X vector and a Z vector;

$$\mathcal{B}(S) = (s_{x_1}, s_{x_2}, ..., s_{x_n}|s_{z_1}, s_{z_2}, ..., s_{z_n}|s) = (\vec{s}_x|\vec{s}_z|s).$$

From this we can reconstruct the stabilizer operator via;

$$S = (-1)^s i^{\vec{s}_x \cdot \vec{s}_z} X(\vec{s}_x) Z(\vec{s}_z) = (-1)^s \bigotimes_i^n i^{s_{x_i} s_{z_i}} X^{s_{x_i}} Z^{s_{z_i}}, \tag{4.19}$$

which is the analogue of eq. (4.17).

From this we can uniquely characterize any stabilizer Pauli operator by $2n$ binary numbers with an additional binary number to denote the phase associated with the operator, as per definition 4.3.3. This then forms a group representation of the projective Pauli group, where multiplication of operators is replaced with addition of binary vectors. This allows us to define the binary-symplectic representation of the stabilizer operators as;

Definition 4.3.7 *The binary-symplectic representation is a map, $\mathcal{B}$, from the stabilizer Pauli operators, $\mathcal{S}_n$, to a binary-symplectic vector space;*

$$\mathcal{B}: \mathcal{S}_n \mapsto \mathbb{Z}_2^{2n} \times \mathbb{Z}_2.$$

We can represent any stabilizer operator, S, by;

$$\mathcal{B}(S) = (\underline{a}, s) \,, \; \forall S \in \mathcal{S}_n,$$

where $\underline{a} = (\underline{a}_x, \underline{a}_z) \in \mathbb{Z}_2^{2n}$ is the vector representing the projective Pauli operator and $s \in \mathbb{Z}_2$ represents the phase associated with it, as per eq. (4.18) and definition 4.3.3.

The binary-symplectic representation in the context of this thesis allows us to index any projective Pauli operators as P_a, where $a \in \mathbb{Z}_2^{2n}$. This binary-symplectic representation can then be extended to characterize stabilizer groups with n binary-symplectic vectors representing the generators of the stabilizer groups.

The commutation relations between two Pauli operators, e.g. P_a and P_b, can be calculated in the binary symplectic representation by the *binary-symplectic inner product*;

$$[a, b] = \frac{1}{2} \left(a_x.b_z + a_z.b_z \right) \; mod \; 2,$$

which can easily be derived by using eq. (4.19) and the fact that $P_a P_b P_a^\dagger = (-1)^{[a,b]} P_b$.

Finally it should be noted that special care must be taken in calculating the phase bit, s, when composing Pauli operators as it is not a straight forward addition, but rather some polynomial function, for more details see section 4.5.

The Gottesman-Knill Theorem

While the Gottesman-Knill Theorem [31], and corresponding improvement by Aaronson and Gottesman [2], are not directly relevant to the work presented here. It does form a key motivation for why we are interested in representing the n-qubit stabilizer formalism with a ψ-epistemic model, because due to these theorems we know that the n-qubit stabilizer formalism represents a subtheory of quantum theory that admits an efficient classical simulation. Therefore, we would expect it to admit a classical model, i.e. one that is not ψ-ontic.

The theorem states that there is a classical simulation algorithm that can simulate any n-qubit stabilizer circuit in polynomial time, $O(n^3)$ for the original theorem [31] and $O(n^2)$ for the improved simulation [2]. To achieve these run-times the theorems use a very particular kind of stabilizer circuit, namely; An input state that is the computational all-zero state $|0\rangle^{\otimes n}$. Clifford gates and Pauli gates from $\{\mathbb{I}, X_i, Y_i, Z_i, H_i, S_i, CNOT_{i,j}\}$. And intermediate computational basis measurements, i.e. the Z basis. This particular set of circuits is equivalent to the full set of Clifford circuits in the stabilizer formalism.

The key thing to extract from the Gottesman-Knill Theorem, for the purposes of this thesis, is the use of stabilizer tableaus. To represent a stabilizer group of some stabilizer state ψ we use a binary tableau, $\mathcal{T}_{\psi,G}$, representing a particular choice of generators for the group, $G = \{g_1, g_2, \ldots, g_n\}$;

$$
\mathcal{T}_{\psi,G} := \left(\begin{array}{cc|c} a_{x,g_1} & a_{z,g_1} & s_{g_1} \\ a_{x,g_2} & a_{z,g_2} & s_{g_2} \\ \vdots & \vdots & \vdots \\ a_{x,g_n} & a_{z,g_n} & s_{g_n} \end{array} \right). \tag{4.20}
$$

Operations in the simulation algorithm are then performed on this tableau. Note, the improved simulation algorithm adds n extra rows representing a generating set for the full Pauli group, termed *destabilizer generators*, allowing for the improvement in simulation runtime.

However, the computational model of the stabilizer formalism given by these tableaus constitutes a ψ-ontic model of the stabilizer formalism, where ψ-onticity is often regarded as a non-classical property. We can see this by looking at the ontological elements as defined by the Gottesman-Knill theorem. The ontology of the Gottesman-Knill algorithm is the set of all tableaus $\Lambda = \{\mathcal{T}_{\psi,G}\}$. Epistemic states are then delta-functions over a particular tableau, making the model ψ-ontic. However there is already a glaring issue with this, there are as many tableaus representing a state as there are choices of generators for that state's stabilizer group. This makes makes the model highly preparation contextual, as even pure states are contextually represented. Therefore this begs the question: Is there a more natural ontological model of the n-qubit stabilizer formalism? This question motivated the work presented in this chapter, and lead to the ψ-epistemic model presented in chapter 5, and as a consequence the results in chapter 6.

4.3.3   Quasi-Probability Representations and Frames

In this chapter we will be focused on the n-qubit discrete Wigner function, which is an example of a *quasi-probability representation* (QPR) of quantum theory. However unlike previous definitions of the Wigner function in this work we will allow the representation to switch between *different* Wigner functions, i.e. switching frames - the frame-switching Wigner function. To perform this mathematical trickery it will be useful to cover some of the basic concepts from the formalism of quasi-probability representations.

The most fundamental definition of a quasi-probability representation is that given by Spekkens [78]. Every quantum density operator ρ on a Hilbert space $\mathcal{H}$ is associated with a

real-valued measure μ_ρ over some measurable space Λ, such that $\int_\Lambda \mu_\rho d\lambda = 1$. Note as this function is only constrained to be real valued it can be negative. Hence the term *quasi-probability* representation, as these negative values are thought of in the same context as negative probabilities.

Every linear CP map Φ in the space of linear operators on $\mathcal{H}$ is associated to a real-valued measurable function $\Gamma_\Phi(\lambda', \lambda)$ over the measurable space $\Lambda \times \Lambda$, such that $\int_\Lambda \Gamma_\Phi(\lambda', \lambda) = 1, \forall \lambda \in \Lambda$. Here we have been slightly sloppy as this definition could allow for switching between representations, by adjusting the definition of the image of the map. Indeed, it is possible to define a pure "representation switching" operation, and any quantum map can be thought of as having been performed before or after the switch of representation.

Every quantum measurement, POVM, $\{E_k\}$ on $\mathcal{H}$ is associated to a set of real-valued measurable functions $\{\xi_k\}$ over Λ, such that $\sum_k \xi_k(\lambda) = 1, \forall \lambda \in \Lambda$. With the trivial measurement $\mathbb{I}$ being represented by the unit function over Λ.

Putting all these definitions together, a quasi-probability representation is a faithful representation if it reproduces the quantum statistics, i.e. it satisfies;

$$\mathrm{Tr}(E_k \Phi(\rho)) = \int_\Lambda \int_\Lambda \xi_k(\lambda') \Gamma_\Phi(\lambda', \lambda) \mu_\rho(\lambda) d\lambda d\lambda'. \tag{4.21}$$

There are several things to note about this definition of a quasi-probability representation. Firstly, the notions of generalized non-contextuality are essentially baked into the definitions. With every quantum operation having a one-to-one correspondence with an element of the quasi-probability representation. Secondly, the definitions are effectively the same as the definitions for an ontological model, without the requirement of positivity and the one-to-one correspondence between operational elements and representations. Therefore we can see that if we impose positivity on the definitions of a quasi-probability representation then we automatically get the definitions of a non-contextual ontological model, and similarly in the reverse direction. Lastly, the definitions of a quasi-probability representation are actually a subset of the super-operator representations, for example the Pauli-Louville representation, where we impose normalization on the representation. This in my eyes casts doubt on whether it is correct to interpret negative values as negative probabilities. In the definitions of the Wigner function I give an argument proposing these negative values are better seen as expectation value of outcomes, rather than mystical negative probabilities.

The above definitions provide a very abstract way of thinking about quasi-probability representations. A more accessible approach is that given by the frame formalism, where

more detail can be found in [27]. As before, a frame is defined by referring to some measurable space Λ, finite or infinite, referred to as a phase-space. However unlike the abstract definition of a QPR, to each point in the phase-space we associate two Hermitian operators F_λ and D_λ. The set F_λ, $\mathcal{F} = \{F_\lambda | \, F_\lambda \in \mathbb{H}(\mathcal{H})\}$, is called the *primal frame*, and the set of D_λ, $\mathcal{D} = \{D_\lambda | \, D_\lambda \in \mathbb{H}(\mathcal{H})\}$, is called the *dual frame*. The primal and dual frames must satisfy;

$$A = \sum_\lambda \langle F_\lambda, A \rangle \, D_\lambda, \tag{4.22}$$

given some inner product $\langle \cdot, \cdot \rangle \in \mathbb{R}$ on the space and the above equation holds for all $A \in \mathbb{H}(\mathcal{H})$ we wish to represent in the frame representation. We refer to eq. (4.22) as the *recombination formula*. Note, for eq. (4.22) to hold we only require span($\mathcal{F}$) and span($\mathcal{D}$) only include the operators we are trying to represent. Typically it is the case that span($\mathcal{F}$) = $\mathbb{H}(\mathcal{H})$ and span($\mathcal{D}$) = $\mathbb{H}(\mathcal{H})$, but these conditions are not required. For example the frames used in the Delfosse et al. rebit Wigner function [25] has frames that do not satisfy this property, yet are properly defined frame representations. As we will be working with Wigner functions where the set of phase-point operators do span the space of Hermitian operators, I will assume that a primal and dual frame pair span the space of Hermitian operators, unless stated otherwise.

As density operators are Hermitian operators we can apply the recombination formula to them;

$$\rho = \sum_\lambda \langle F_\lambda, \rho \rangle \, D_\lambda. \tag{4.23}$$

Given that $\langle F_\lambda, \rho \rangle \in \mathbb{R}$ we define the quasi-probability representation of a state to be $\mu_\rho(\lambda) = \langle F_\lambda, \rho \rangle$, i.e. a state is represented in the *primal frame*. Conversely POVMs are represented in the *dual frame*, however to reproduce the quantum statistics here we must use the Hilbert-Schmidt inner product. So every POVM element E_k is represented by the measurable function $\xi_k(\lambda) = \mathrm{Tr}(E_k D_\lambda)$, which we often refer to as *response functions*. From this we can see the frame formalism reproduces the quantum statistics;

$$\mathrm{Tr}(\rho E_k) = \mathrm{Tr}(\sum_\lambda \langle F_\lambda, \rho \rangle \, D_\lambda E_k),$$

$$= \sum_\lambda \langle F_\lambda, \rho \rangle \, \mathrm{Tr}(E_k D_\lambda),$$

$$= \sum_\lambda \mu_\rho(\lambda) \xi_k(\lambda) = \mu_\rho . \xi_k.$$

Further, imposing the normalization constraints of a QPR on the frame representation constrains the structure of $\mathcal{F}$ and $\mathcal{D}$. From the normalization of a state we have;

$$\sum_\lambda \mu_\rho(\lambda) = \sum_\lambda \langle F_\lambda, \rho \rangle,$$

$$= \left\langle \sum_\lambda F_\lambda, \rho \right\rangle = 1.$$

Therefore we have $\sum_\lambda F_\lambda = \mathbb{I}^9$. From the requirement that $\xi_\mathbb{I}(\lambda) = 1$, $\forall \lambda$ we have $\mathrm{Tr}(D_\lambda) = 1$, $\forall D_\lambda \in \mathcal{D}$.

To represent quantum CP maps we again turn to the recombination formula. Given some channel Φ we denote the output of the channel as $\rho' = \Phi(\rho)$. Assuming both ρ and ρ' are represented by the frame QPR we have;

$$\rho = \sum_\lambda \langle F_\lambda, \rho \rangle D_\lambda,$$

$$\Phi(\rho) = \sum_\lambda \langle F_\lambda, \rho \rangle \Phi(D_\lambda),$$

$$\rho' = \sum_\lambda \langle F_\lambda, \rho' \rangle D_\lambda = \sum_\lambda \langle F_\lambda, \Phi(\rho) \rangle D_\lambda.$$

Expanding $\Phi(D_\lambda) = \sum_{\lambda'} \langle F_{\lambda'}, \Phi(D_\lambda) \rangle D_{\lambda'}$ using the recombination formula and equating lines 2 and 3 of the above equations gives;

$$\sum_\lambda \langle F_\lambda, \rho' \rangle = \sum_{\lambda, \lambda'} \langle F_{\lambda'}, \rho \rangle \langle F_\lambda, \Phi(D_{\lambda'}) \rangle D_\lambda. \tag{4.24}$$

Therefore, as the coefficients for each D_λ must be equal we can see that $\langle F_\lambda, \Phi(\rho) \rangle = \sum_{\lambda'} \langle F_{\lambda'}, \rho \rangle \langle F_\lambda, \Phi(D_{\lambda'}) \rangle$. From this we define $\Gamma_\Phi(\lambda, \lambda') = \langle F_\lambda, \Phi(D_{\lambda'}) \rangle$ to be the transition matrix representing Φ in the QPR. We call this transition map *quasi-stochastic* as it inherits normalization from the normalization of the primal frame; $\sum_\lambda \langle F_\lambda, \Phi(D_{\lambda'}) \rangle = \langle \mathbb{I}, \Phi(D_{\lambda'}) \rangle = 1$, $\forall D_{\lambda'} \in \mathcal{D}$, or in terms of the QPR language $\sum_\lambda \Gamma_\Phi(\lambda, \lambda') = 1$, $\forall \lambda \in \Lambda$.

We can apply a similar technique as we did to the representation of quantum maps to define how we switch from one frame to another. Suppose we wished to switch our representation from the primal/dual frame pair $\mathcal{F}$ and $\mathcal{D}$ to another primal/dual frame

---

[9]Or the projector onto the space $\mathcal{F}$ spans.

pair $\mathcal{G}$ and $\mathcal{H}$, by applying the recombination formula we have;

$$\rho = \sum_\lambda \langle F_\lambda, \rho \rangle \, D_\lambda,$$

$$= \sum_{\lambda, \lambda'} \langle F_\lambda, \rho \rangle \, \langle G_{\lambda'}, D_\lambda \rangle \, H_{\lambda'},$$

$$= \sum_{\lambda'} \langle G_{\lambda'}, \rho \rangle \, H_{\lambda'}.$$

Therefore again it must be the case that $\langle G_{\lambda'}, \rho \rangle = \sum_\lambda \langle F_\lambda, \rho \rangle \, \langle G_{\lambda'}, D_\lambda \rangle$. From this we identify $\Gamma_{\{\mathcal{F},\mathcal{D}\} \to \{\mathcal{G},\mathcal{H}\}}(\lambda, \lambda') = \langle G_{\lambda'}, D_\lambda \rangle$ as the transition map between two different representations.

As we shall see in the following sections the transition map between two frames can be negatively represented, but can still used to define stochastic maps between two frames. As long as we compose the map with a quantum operation and limit the set of input states and output states. However, this allows in the possibility of transformation contextuality. In much the same way as the work presented in chapter 2: Wherein the 8-state *Wigner function* has negatively represented transition matrices for the single-qubit Clifford gates, but there still exists a sensible way to define stochastic transition maps for the Clifford gates, which define the 8-state *model*[10]. Finally there is some confusion within our academic group as to whether the frame formalism is actually equivalent to the definitions of a quasi-probability representation given by Spekkens. However, investigating this question is not relevant to the work in this thesis, but does pose an interesting theoretical question.

4.3.4   The Discrete Wigner Function

The Wigner function is a method by which we can represent quantum states while retaining some notion of a classical phase-space. As mentioned previously Wigner functions are quasi-probability representations of quantum states. Here our interest lies in the n-qubit stabilizer formalism, and therefore we will be working with Wigner functions defined on the Hilbert space $\mathcal{H}_{2^n}$.

In the Wigner function representation we use frames that form operator bases for the set of all Hermitian operators on $\mathcal{H}$, $\mathbb{H}(\mathcal{H}_d)$. So $\mathcal{F} = \{A_a\}$, $\mathrm{Tr}(A_a A_b) = \delta_{a,b}$, and $span(\{A_a\}) = \mathbb{H}(\mathcal{H}_d)$ therefore $|\{A_a\}| = d^2$. Further, for the Wigner functions we are interested in, these

---

[10]Here I have used "8-state Wigner function" and "8-state model" to draw attention to the fact that these two models are only equivalent if transformations are not included in the operational theory.

operator bases are self dual, i.e. $\mathcal{F} = C\mathcal{D}$, where C is some normalization constant. The choice of inner product used in the Wigner function is the standard Hilbert-Schmidt inner product, i.e. $\langle A, B \rangle = \text{Tr}(AB)$.

As we are interested in the the n-qubit stabilizer subtheory we limit ourselves to Wigner functions defined over $\mathcal{H}_{2^n}$. This leaves us with the question: which set of operators do we use for $\{A_a\}$? Ideally we wish to use operators that give positive representations of stabilizer states. We'd also like the set of frame operators to generate from an *origin* frame operator, much like in classical phase-space representations. A natural choice is to follow the definitions of the qudit Wigner function, thus we arrive at our first attempt at defining a n-qubit stabilizer Wigner function's phase-point operators $\mathcal{F} = \{A_a\}$;

$$A_0 = \frac{1}{d^2} \sum_{P_b \in \widetilde{\mathcal{P}}_n} P_b, \tag{4.25}$$

$$A_a = P_a A_0 P_a = \frac{1}{d^2} \sum_{P_b \in \widetilde{\mathcal{P}}_n} (-1)^{[a,b]} P_b, \ P_a \in \widetilde{\mathcal{P}}_n, \tag{4.26}$$

where A_0 is the origin phase-point operator and the P_a the *translation* operators. This definition is the analogue of the widely used qudit Wigner function [29]. However for qubits[11] we require a more general definition if we are to have all stabilizer states positively represented, under some choice of Wigner function. For the moment let us use the phase-point operators given by eq. (4.25) to define one example of a discrete Wigner function for qubits.

The representation of a quantum state in the Wigner function is given by;

$$\mu_\rho(\lambda) = W_a(\rho) = \text{Tr}\left(\text{A}_\text{a}\rho\right). \tag{4.27}$$

Further as the Wigner function frame is self-dual we also have;

$$\xi_a(E_k) = \text{Tr}\left(C\text{A}_\text{a}\text{E}_\text{k}\right). \tag{4.28}$$

Finally, transformations in the Wigner function are represented by;

$$\Gamma_\Phi(a', a) = \text{Tr}\left(\text{A}_\text{a}\Phi(C\text{A}_{\text{a}'})\right). \tag{4.29}$$

However despite this being the canonical definition of a transition map, which crucially is a linear mapping, this does not mean this definition is well applied or even used in the

---

[11]And qudits, although the additional phases turn out to be trivial.

literature. In fact far from it, it is often the case that something like the following will be seen;

$$\Phi(\rho) = \sum_a \text{Tr}\left(A_a\rho\right)\Phi(A_a) = \sum_a \text{Tr}\left(A_a\rho\right)A_{\phi(a)}, \tag{4.30}$$

where ϕ is taken to be an bijective function on the set of a, i.e. it permutes a in some manner[12]. If such a ϕ can be found, for some set of transformations $\{\Phi\}$, it is then argued that this set of transformations can be represented by a corresponding permutation, ϕ, of the phase space. However this is not strictly true, in the QPR sense, an example of which lead to the result presented in chapter 2, where as for the qudit Wigner function it turns out that eq. (4.29) and eq. (4.30) *are* equivalent.

For the following work it is useful to see the consequences of this non-equivalence in representations of transformations. To do so we can use the 8-state Wigner function and Model as examples. For the 8-state model it *is* the case that;

$$CA^p_{(x,z)}C^\dagger = \frac{1}{8}C\left(\mathbb{I} + (-1)^x X + (-1)^{x+z+p}Y + (-1)^z Z\right)C^\dagger,$$

$$= \frac{1}{8}\left(\mathbb{I} + (-1)^{c(x)}X + (-1)^{c(x+z)+p'}Y + (-1)^{c(z)}Z\right) = A^{p'}_{(c(x),(c(z))}, \ \forall C \in Cl_2,$$

where $c(x)$, and $c(z)$ represents the permutation the Clifford acts on the X and Z phase. For example, the Hadamard gate, H, gate action on a phase-point operator is given by;

$$HA^p_{(x,z)}H^\dagger = \frac{1}{8}H\left(\mathbb{I} + (-1)^x X + (-1)^{x+z+p}Y + (-1)^z Z\right)H^\dagger,$$

$$= \frac{1}{8}\left(\mathbb{I} + (-1)^z X + (-1)^{x+z+p+1}Y + (-1)^x Z\right) = A^{p+1}_{(z,x)}.$$

Therefore, the Hadamard gate can be represented by the permutation $H : (x, z, p) \mapsto (z, x, p + 1)$. However when we look at the representation of the the Hadamard using eq. (4.29) we find it is negatively represented;

$$4\Gamma_H = \begin{pmatrix}
1 & 1 & -1 & 1 & 2 & 0 & 0 & 0 \\
1 & -1 & 1 & 1 & 0 & 0 & 0 & 2 \\
-1 & 1 & 1 & 1 & 0 & 0 & 2 & 0 \\
1 & 1 & 1 & -1 & 0 & 2 & 0 & 0 \\
2 & 0 & 0 & 0 & 1 & 1 & -1 & 1 \\
0 & 0 & 0 & 2 & 1 & -1 & 1 & 1 \\
0 & 0 & 2 & 0 & -1 & 1 & 1 & 1 \\
0 & 2 & 0 & 0 & 1 & 1 & 1 & -1
\end{pmatrix}.$$

---

[12]Often referred to as *covariance*.

Indeed *all* Clifford gates in the 8-state Wigner function are negatively represented, including the identity gate. This mismatch between eq. (4.29) and eq. (4.30) is what allows for the emergence of transformation contextuality, as eq. (4.29) is a linear map where as eq. (4.30) is not.

4.4   The Frame-Switching Wigner Function

The n-qubit Wigner Function phase point operators given in the previous section, eq. (4.25) are not the only choice of phase-point operators we can use to define a n-qubit Wigner function. Indeed to have any hope of representing the n-qubit stabilizer formalism positively we must turn to other options. For example, the representations of all Bell states, e.g. $|B_{00}\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$, using the phase-point operators given in eq. (4.25) are negative.

The way we will achieve this is to define a class of phase-point operators based on what we term the *phase function*;

> **Definition 4.4.1** *The **phase function**, γ, is a mapping from the projective Pauli operators, $\widetilde{\mathcal{P}}_n$, to $\mathbb{Z}_2$;*
>
> $$\gamma : \widetilde{\mathcal{P}}_n \mapsto \mathbb{Z}_2, \tag{4.31}$$
>
> *such that $\gamma(\mathbb{I}) = 0$. We define the set of n-qubit phase functions to be $\gamma_n = \left\{ \gamma | \widetilde{\mathcal{P}}_n \mapsto \mathbb{Z}_2 \right\}$*

Using the phase function we then define the origin of a Wigner function, indexed by γ to be;

$$A_0^\gamma = \frac{1}{4^n} \sum_{P_b \in \widetilde{\mathcal{P}}_n} (-1)^{\gamma(b)} P_b$$

The other phase-point operators are defined as they were previously, using translation operators;

$$A_a^\gamma = P_a A_0^\gamma P_a \frac{1}{4^n} \sum_{P_b \in \widetilde{\mathcal{P}}_n} (-1)^{\gamma(b)+[a,b]} P_b.$$

Therefore let us define the frame associated to a phase function γ as;

$$\mathcal{F}^\gamma = \{A_a^\gamma\}_a = \left\{ P_a A_0^\gamma P_a^\dagger | P_a \in \widetilde{\mathcal{P}}_n \right\}. \tag{4.32}$$

Here I note that the above definition of the origin can be improved. However, it is not technically necessary for the results present here and complicates the frame switching procedure. To see this consider that a perfectly acceptable phase function is given by $\gamma(b) = [c, b]$, $c \in \mathcal{B}(\widetilde{\mathcal{P}}_n)$. I.e. given this definition of γ we can chose *any* phase-point operator from eq. (4.25) as the origin. To avoid this "double counting"[13] we can define an equivalence class of phase functions $\mathcal{C}_\gamma = \left\{ \gamma' \mid \gamma' = \gamma + [\cdot, c]; \ c \in \mathcal{B}(\widetilde{\mathcal{P}}_n) \right\}$. Unfortunately defining these equivalence classes does not simplify our problem due to the symmetry of the classes, which makes any choice of class representative a purely aesthetic one.

Now we must prove that any choice of γ gives a valid Wigner function. To do so we must check that $\mathcal{F}^\gamma$ satisfies the frame conditions $\sum_\lambda F_\lambda = \mathbb{I}$ and $\mathrm{Tr}\,(D_\lambda) = 1$, where Hermiticity of the phase-point operators follows naturally from the Hermicitiy of the Pauli operators. The condition that the primal frame sums to the identity is easy to derive;

$$
\begin{aligned}
\sum_\lambda F_\lambda &= \sum_a A_a^\gamma, \\
&= \frac{1}{4^n} \sum_a \sum_b (-1)^{\gamma(b) + [a,b]} P_b, \\
&= \frac{1}{4^n} \sum_b (-1)^{\gamma(b)} \left[\sum_a (-1)^{[a,b]} \right] P_b = \mathbb{I}.
\end{aligned}
$$

For the last line we have used the definition that $\gamma(\mathbb{I}) = 0$, $\forall \gamma$ and that $\sum_a (-1)^{[a,b]} = 4^n$ if and only if $b = \mathcal{B}(\mathbb{I})$ and is equal to zero otherwise, as $[\cdot, b]$ is a balanced vector[14].

Next we must check is that $\mathcal{F}^\gamma$ forms a self-dual frame;

Lemma 4.4.1 *Any Wigner function frame* $\mathcal{F}^\gamma = \left\{ P_a A_0^\gamma P_a^\dagger \mid P_a \in \widetilde{\mathcal{P}}_n \right\}$ *forms a self-dual frame for all choices of* γ. *I.e. it satisfies;*

$$
A = \frac{1}{C} \sum_a \mathcal{F}^\gamma(a) \mathrm{Tr}\,(\mathcal{F}^\gamma(a) A), \ \forall A \in \mathbb{H}(\mathcal{H}_{2^n}) \tag{4.33}
$$

Where C *is a proportionality constant the relates the dual to the primal frame, i.e.* $\mathcal{D} = \frac{1}{C}\mathcal{F}$.

---

[13]More like 4^n counting.
[14]i.e. it has equal number of zeros and ones.

Proof. Our frame elements are given by;

$$\mathcal{F}(a) = \frac{1}{2^n} \sum_b (-1)^{\gamma(b)+[a,b]} P_b.$$

We can use the projective Pauli operators, e.g. $P_i \in \widetilde{\mathcal{P}}_n$, to show that eq. (4.33) is satisfied for all $\mathcal{F}^\gamma$, as they form a basis for the set of Hermitian operators;

$$\frac{1}{C} \sum_a \mathcal{F}^\gamma(a) \mathrm{Tr}\left(\mathcal{F}^\gamma(a) P_i \right) = \frac{1}{4^{2n}C} \sum_a \left(\sum_b (-1)^{\gamma(b)+[a,b]} P_b \right) \mathrm{Tr}\left(\left(\sum_{b'} (-1)^{\gamma(b')+[a,b']} P_{b'} \right) P_i \right),$$

$$= \frac{1}{4^n 2^n C} \sum_a \sum_b (-1)^{\gamma(b)+[a,b]} P_b (-1)^{\gamma(i)+[a,i]},$$

$$= \frac{1}{4^n 2^n C} \sum_b (-1)^{\gamma(b)+\phi(i)} \left[\sum_a (-1)^{[a,b+i]} \right] P_b = \frac{1}{2^n C} P_i,$$

where the sums are taken over the projective Pauli group and $C = \frac{1}{2^n}$.

Therefore the frame $\mathcal{F}^\gamma$ is self-dual $\forall \gamma$. ∎

We therefore can identify the dual frame for these Wigner functions as $\mathcal{D}^\gamma = 2^n \mathcal{F}^\gamma$. From this it is easy to see that the phase-point operators satisfy the property $\mathrm{Tr}\,(D_\lambda) = \mathrm{Tr}\,(A_a^\gamma) = 1$, $\forall a, \gamma$, as we have $\mathrm{Tr}\,(P) = 0$, $\forall P \in \widetilde{\mathcal{P}}_n \backslash \mathbb{I}$.

Finally all choices of γ additionally lead to $\mathcal{F}^\gamma$ being an operator basis for $\mathbb{H}(\mathcal{H}_{2^n})$, i.e they satisfy $\mathrm{Tr}\,(A_a^\gamma B_b^\gamma) = \delta_{a,b}$, where $A_a^\gamma \in \mathcal{F}^\gamma$ and $B_b^\gamma \in \mathcal{D}^\gamma$;

$$\mathrm{Tr}\,(A_a^\gamma B_b^\gamma) = \frac{1}{4^n 2^n} \mathrm{Tr}\left(\sum_{c,c'} (-1)^{\gamma(c)+\gamma(c')+[a,c]+[b,c']} P_c P_{c'} \right), \tag{4.34}$$

$$= \frac{1}{4^n} \sum_c (-1)^{[a+b,c]} = \delta_{a,b}, \tag{4.35}$$

where we have used the trace-orthonormality of projective Pauli operators, $\mathrm{Tr}\,(P_a P_b) = 2^n \delta_{a,b}$, to go from line 1 to 2. The fact that $(\mathcal{F}^\gamma, \mathcal{D}^\gamma)$ forms a basis is unsurprising given that $|\mathcal{F}^\gamma| = 4^n$, i.e. the number of operators required to define an operator basis for $\mathbb{H}(\mathcal{H}_{2^n})$.

4.4.1 Stabilizer State Representations in the Frame-Switching Wigner Function

With these different classes of Wigner functions established we can now look at how the frame-switching Wigner function works. The idea will be that we model stabilizer states in frames where they are positively represented. Then, when an quantum operation is applied we switch frames to make sure the state stays positively represented. However, as we shall see this is not possible when we try to model measurement update.

The first step in this procedure is to identify which frames a stabilizer state is positively represented in. However prior to doing so we will need a straightforward preliminary lemma;

> **Lemma 4.4.2** *If a pure stabilizer state $\psi \in \mathcal{S}(\mathcal{H}_{2^n})$ is positively represented by a Wigner function W^γ, i.e. $W_a^\gamma(\rho_\psi) \geq 0$, $\forall a \in \mathcal{B}(\widetilde{\mathcal{P}}_n)$, then all states of the form $P\,|\psi\rangle$, $P \in \widetilde{\mathcal{P}}_n$ are also positively represented.*

Proof. We have;

$$W_a^\gamma(P\rho_\psi P) = \mathrm{Tr}\left(\mathrm{A}_a^\gamma \mathrm{P}\rho_\psi \mathrm{P}\right) = \mathrm{Tr}\left(\mathrm{PA}_a^\gamma \mathrm{P}\rho_\psi\right),$$
$$= \mathrm{Tr}\left(\mathrm{A}_{a+p}^\gamma \rho_\psi\right) \geq 0.$$

■

One of the consequences of this lemma is that the bases of the form

$$\left\{ S(\psi) = \left\langle (-1)^{k_1} G_1, (-1)^{k_2} G_2, \ldots, (-1)^{k_n} G_n, \right\rangle \mid \vec{k} \in \mathbb{Z}_2^n \right\}$$

are all positively represented in the same Wigner function. I refer to these bases as the *Pauli bases*, as the set can be constructed via projecting on the joint subspaces of n commuting Pauli operators. Famous examples of such bases are the computational basis, the X basis, the Y basis, and the maximally entangled basis. However note these are not the only orthonormal bases that can be constructed in the stabilizer formalism. For example, the basis;

$$\left\{ |00\rangle, \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle), \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle), |11\rangle \right\},$$

cannot be constructed from a single set of generators, as we have;

$$S(|00\rangle) = \{\mathbb{I}, \mathbb{I}Z, Z\mathbb{I}, ZZ\},$$

$$S(\frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)) = \{\mathbb{I}, XX, -ZZ, YY\},$$

$$S(\frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)) = \{\mathbb{I}, -XX, -ZZ, -YY\},$$

$$S(|11\rangle) = \{\mathbb{I}, -\mathbb{I}Z, -Z\mathbb{I}, ZZ\}.$$

Using lemma 4.4.2 we can prove our main stabilizer state positivity theorem;

Theorem 4.4.3 *A pure stabilizer state $\psi \in \mathcal{S}(\mathcal{H}_{2^n})$ is positively represented, $W_a^\gamma(\rho_\psi) \geq 0$, $\forall a \in \mathcal{B}(\widetilde{\mathcal{P}}_n)$, if and only if the phase function γ is of the form;*

$$\gamma(b) = s_b + [c, b], \ \forall b \in \mathcal{B}(S(\psi)), \ c \in \mathcal{B}(\widetilde{\mathcal{P}}_n) \tag{4.36}$$

where s_b is defined as the phase associated to the Stabilizer group elements;

$$S_b = (-1)^{s_b} P_b, \ \forall b \in \mathcal{B}(S(\rho)) \text{ and } P_b \in \widetilde{\mathcal{P}}_n.$$

Proof. **Forward direction:**

$$\gamma(b) = s_b + [c, b], \ \forall b \in \mathcal{B}(S(\psi)), \ c \in \mathcal{B}(\widetilde{\mathcal{P}}_n) \implies W_a^\gamma(\rho_\psi) \geq 0, \ \forall a \in \mathcal{B}(\widetilde{\mathcal{P}}_n).$$

We can express the state ρ_ψ as;

$$\rho_\psi = \frac{1}{2^n} \sum_{S \in S(\psi)} S = \frac{1}{2^n} \sum_{b \in \mathcal{B}(S(\psi))} (-1)^{s_b} P_b.$$

Therefore the Wigner function of this state with our choice of phase function γ is given by;

$$W_a^\gamma(\rho_\psi) = \text{Tr}\left(\frac{1}{4^n}\left[\sum_{b \in \mathcal{B}(\widetilde{\mathcal{P}}_n)} (-1)^{\gamma(b)+[a,b]} P_b\right] \frac{1}{2^n}\left[\sum_{b' \in \mathcal{B}(S(\psi))} (-1)^{s_{b'}} P_{b'}\right]\right).$$

Using trace-orthonormality of the projective Pauli operators, $\text{Tr}(P_b P_a) = 2^n \delta_{b,a}$, and applying our assumption, $\gamma(b) = s_b + [c, b]$, $\forall b \in \mathcal{B}(S(\psi))$, $c \in \mathcal{B}(\widetilde{\mathcal{P}}_n)$, we have;

81

$$W_a^\gamma(\rho_\psi) = \frac{1}{2^{2n}} \sum_{b \in \mathcal{B}(\mathcal{S}(\psi))} (-1)^{\gamma(b)+s_b+[a,b]} = \frac{1}{2^{2n}} \sum_{b \in \mathcal{B}(\rho)} (-1)^{[a',b]},$$

where we have used the linearity of the inner product, and set $a' = a + c$, $c \in \mathbb{Z}_2^{2n}$.

As $\mathcal{B}(\mathcal{S}(\psi))$ forms a group we can add any element of the group to all members of the group and still have the same set of elements, therefore;

$$\frac{1}{2^{2n}} \sum_{b \in \mathcal{B}(\mathcal{S}(\psi))} (-1)^{[a',b]} = \frac{1}{2^{2n}} \sum_{b \in \mathcal{B}(\mathcal{S}(\psi))} (-1)^{[a',b+d]},$$

$$= \frac{1}{2^{2n}} (-1)^{[a',d]} \sum_{b \in \mathcal{B}(\mathcal{S}(\psi))} (-1)^{[a',b]}.$$

Labelling $\sum_{b \in \mathcal{B}(\mathcal{S}(\psi))} (-1)^{[a',b]} = S$, we have $S = (-1)^{[a',d]} S$. If $\exists d \in \mathcal{B}(\mathcal{S}(\psi))$ such that $[a',d] = 1$ then $S = -S \implies S = 0$. If no such d exists $[a',d] = 0$, $\forall d \in \mathcal{B}(\mathcal{S}(\psi))$ and therefore $S = 2^n$. Considering both cases implies that $W_a^\phi(\rho) \geq 0$, $\forall a$.

Backward direction:

$$W_a^\gamma(\rho_\psi) \geq 0, \, \forall a \in \mathcal{B}(\widetilde{\mathcal{P}}_n) \implies \gamma(b) = s_b + [c,b], \, \forall b \in \mathcal{B}(S(\psi)), \, c \in \mathcal{B}(\widetilde{\mathcal{P}}_n).$$

To proceed with the proof we will actually step outside the definitions of QPRs and use a result from non-contextuality. Namely that in any preparation non-contextual ontological model of quantum theory of dimension d we have;

$$\mu_\psi(\lambda) = d\mu_{\mathbb{I}/d}(\lambda), \, \forall \mu_\psi, \, \forall \lambda \in \text{supp}(\mu_\psi).$$

Given that there exists an orthonormal basis ψ is an element of, as proven in the introduction of this thesis[15]. If the stabilizer state ψ is positively represented by a Wigner function W^γ, then W^γ forms a preparation non-contextual ontological model for the Pauli basis ψ is an element of. As by lemma 4.4.2 if $W_a^\gamma(\rho_\psi) \geq 0$, $\forall a$ then $W_a^\gamma(P\rho_\psi P) \geq 0$, $\forall a$ and $\forall P \in \widetilde{\mathcal{P}}_n$, and $\{P|\psi\rangle\}_{P \in \widetilde{\mathcal{P}}_n}$ forms an orthonormal basis for $\mathcal{H}_{2^n}$.

Therefore we have;

$$W_a^\gamma(\rho_\psi) = 2^n W_a^\gamma(\frac{1}{2^n}\mathbb{I}), \, \forall a \text{ such that } W_a^\gamma(\rho_\psi) > 0. \tag{4.37}$$

---

[15]This fact can be easily proven for postively represented quasi-probability representation as well.

It is a straight forward calculation to see that $W_a^\gamma(\frac{1}{2^n}\mathbb{I}) = 1/4^n$, $\forall a$, which implies $W_a^\gamma(\rho_\psi) = 1/2^n$, $\forall a$ such that $W_a^\gamma(\rho_\psi) > 0$.

Applying the definition of the Wigner function and using our assumption we have;

$$W_a^\gamma(\rho_\psi) = \frac{1}{2^{2n}} \sum_{b \in \mathcal{B}(\mathcal{S}(\psi))} (-1)^{\gamma(b)+s_b+[a,b]} \geq 0, \ \forall a \in \mathcal{B}(\widetilde{\mathcal{P}}_n).$$

Using eq. (4.37) and canceling some factors we can see that;

$$\sum_{b \in \mathcal{B}(\mathcal{S}(\psi))} (-1)^{\gamma(b)+s_b+[a,b]} = \begin{cases} 2^n, \\ 0. \end{cases}$$

As $|\mathcal{B}(\mathcal{S}(\psi))| = 2^n$ the vector $\gamma + s + [a, \cdot]$ must be balanced or constant. Therefore $\gamma(b) + s_b + [a,b] = [c,b]$, $\forall a$ and for some choice of c, implying $\gamma(b) = s_b + [c',b]$, $\forall b \in \mathcal{B}(S(\psi))$, $c \in \mathcal{B}(\widetilde{\mathcal{P}}_n)$. ∎

In theorem 4.4.3 we can see the equivalence classes of phase functions. $\mathcal{C}_\gamma$, appearing again. The binary balanced vector in eq. (4.36) is exactly the definition we used for the equivalence classes. Therefore we can see that if $W_a^\gamma(\rho) \geq 0$, $\forall a$ then $W_a^{\gamma'}(\rho) \geq 0$, $\forall a$, $\forall \gamma' \in \mathcal{C}_\gamma$.

By investigating theorem 4.4.3 we can see that a stabilizer state is positively represented if and only if the phase function of a particular Pauli is equal to the phases on the corresponding Pauli in its stabilizer group, up to some constant or balanced vector. Note this means that positivity is not conditional on Paulis outside the stabilizer group of a state. This fact will be leveraged in chapter 5 when we define a ψ-epistemic ontological model of the stabilizer formalism, which uses a modified version of the phase function. The key foreshadowing I will do here is that by the Wigner functions construction γ provides a value assignment to all Pauli operators.

4.4.2 Clifford Operation Representations in the Frame-Switching Wigner Function

Prior to investigating how frame switching works through a Clifford transformation it will be useful to look at how switching between Wigner functions, with different phase functions,

works in the absence of transformations. From the definition of a frame switching quasi-stochastic map we have;

$$\Gamma_{\gamma_1 \to \gamma_2}(a', a) = \mathrm{Tr}\left(A_{a'}^{\gamma_1} B_a^{\gamma_2}\right),$$
$$\propto \sum_b (-1)^{\gamma_1(b) + \gamma_2(b) + [a+a', b]}. \tag{4.38}$$

The first thing to note is that switching frames between two frames in the same equivalency class, i.e. $\gamma_1, \gamma_2 \in \mathcal{C}_\gamma$, is represented positively. To see this we note that for two phase functions in the same equivalency class we have $A_a^{\gamma_1} = P_c A_a^{\gamma_2} P_c = A_{a+c}^{\gamma_2}$, for some P_c. This means that switching between these two frames is the equivalent to applying a Pauli operator to the system, and under any choice of γ the transition map Γ_P^γ is positively represented;

$$\Gamma_P^\gamma(a', a) = \mathrm{Tr}\left(A_{a'}^\gamma P B_a^\gamma P\right),$$
$$= \mathrm{Tr}\left(A_{a'}^\gamma B_{a+p}^\gamma\right) = \delta_{a', a+p} \geq 0, \ \forall a', a.$$

However, for generic γ_1 and γ_2 eq. (4.38) has no guarantee of being positively represented, in fact it is most likely always negatively represented. This fact eliminates a straightforward frame switching procedure, but we can be clever about how we switch frames when applying a Clifford gate. We will only allow frame switching between particular frames. This leads us a lemma that tells us how Clifford gates allow us to deduce some facts about the positivity of stabilizer states;

Lemma 4.4.4 *If $\rho_{\psi'} = C\rho_\psi C^\dagger$, where ψ and ψ' are stabilizer states, and $W_a^\gamma(\rho_\psi) \geq 0, \ \forall a$, then $W_a^{\gamma'}(\rho_{\psi'}) \geq 0, \ \forall a$.*

Where $C \in Cl_{2^n}$, and $\gamma'(b) = \gamma(C^{-1}(b)) + \gamma_C(b)$, with $\gamma_C(b), C^{-1}(b) : C P_{C^{-1}(b)} C^\dagger = (-1)^{\gamma_C(b)} P_b$.

Proof. Using the definition of the Wigner function, inserting identity and using the cyclic property of the trace we can write;

$$W_a^\gamma(\rho) = \frac{1}{4^n} \mathrm{Tr}\left(A_a^\gamma \rho\right) = \frac{1}{4^n} \mathrm{Tr}\left(C^\dagger C A_a^\gamma C^\dagger C \rho\right),$$
$$= \frac{1}{2^n} \mathrm{Tr}\left(C A_a^\gamma C^\dagger C \rho C^\dagger\right) = \frac{1}{2^n} \mathrm{Tr}\left(C A_a^\gamma C^\dagger \rho'\right).$$

We can then update phase point operators as follows;

$$CA_a^\gamma C^\dagger = CP_a A_0^\gamma P_a^\dagger C^\dagger,$$

$$= P_{a'} CA_0^\gamma C^\dagger P_{a'}^\dagger = P_{a'}\left[\frac{1}{4^n}\sum_b (-1)^{\gamma(b)} CP_b C^\dagger\right] P_{a'}^\dagger,$$

$$= P_{a'}\left[\frac{1}{4^n}\sum_b (-1)^{\gamma(C^{-1}(b))+\gamma_C(b)} P_b\right] P_{a'}^\dagger = A_{a'}^{\gamma'},$$

where $\gamma'(b) = \gamma(C^{-1}(b)) + \gamma_C(b)$, with $\gamma_C := CP_b'C^\dagger = (-1)^{\gamma_C(b)} P_b$, and we have used the definition of the Clifford group to permute through the Clifford operations, as per definition 4.3.4.

Therefore as a Clifford operation is a group automorphism on the Stabilizer operations we have that if $W_a^\gamma(\rho) \geq 0$, $\forall a$, then $W_a^{\gamma'}(\rho') \geq 0$, $\forall a$, with appropriate relabeling. $\blacksquare$

This is very useful for us, it tells us that if we apply a Clifford gate to ψ, which is positively represented by W^γ, then $C\rho_\psi C^\dagger$ is positively represented by $W^{\gamma'}$. All that is left for us to show is that the quasi-probability map representing the Clifford gate and the transition between γ and γ' is positive;

Theorem 4.4.5 *The transition map* $\Gamma_{C,\gamma\to\gamma'}(a', a)$, *where* $\gamma'(b) = \gamma(C^{-1}(b)) + \gamma_C(b)$, *is positive for all* a, a'. *i.e. mathematically;*

$$\Gamma_{C,\gamma\to\gamma'}(a', a) \geq 0, \forall a, a'. \tag{4.39}$$

Proof. From the definition of the transition map we have;

$$\Gamma_{C,\gamma\to\gamma'}(a', a) = \mathrm{Tr}\left(A_{a'}^{\gamma'} CB_a^\gamma C^\dagger\right),$$

$$\propto \mathrm{Tr}\left(\sum_{b'}(-1)^{\gamma'(b')+[a',b']} P_{b'} \sum_b (-1)^{\gamma(C^{-1}(b))+[a,C^{-1}(b)]+\gamma_C(b)} P_b\right),$$

$$= \sum_b (-1)^{\gamma'(b)+\gamma(C^{-1}(b))+\gamma_C(b)+[a',b]+[a,C^{-1}(b)]},$$

$$= \sum_b (-1)^{[a',b]+[a,C^{-1}(b)]},$$

where to go from line 3 to 4 we have used the definition of γ'. Using the fact that $[a, C^{-1}(b)] = [C(a), b]$ we therefore have;

$$\Gamma_{C,\gamma\to\gamma'}(a', a) \propto \sum_b (-1)^{[a',b]+[a,C^{-1}(b)]} = \sum_b (-1)^{[a'+C(a),b]} \geq 0, \forall a, a'.$$

∎

This means that while directly mapping from γ to γ' may be negative and in general the transition matrix $\Gamma_{C,\gamma\to\gamma}$ may also be negative, if we map from γ to γ' in conjunction with applying a Clifford C we will always get a positively represented map, i.e. $\Gamma_{C,\gamma\to\gamma'}$.

So far we have statements about which stabilizer states are positively represented and how applying Clifford transformations maps between these frames, and preserves positivity. Our next element in the frame-switching Wigner function is measurements, which as we shall see pose far more challenging problems.

4.4.3 Stabilizer Measurement Representations in the Frame-Switching Wigner Function

Before investigating how the frame-switching Wigner function attempts to model measurement update we first note that representing how measurement outcomes of Pauli observables are represented;

> **Theorem 4.4.6** *Given a set of phase-point operators defining the dual-frame, $\{B_a^\gamma\}$, of a Wigner function, W^γ.*
>
> *All projectors of the form $\Pi_{s,m} = \frac{1}{2}\left(\mathbb{I} + (-1)^s P_m\right)$, with $s \in \mathbb{Z}_2$ and $P_m \in \widetilde{\mathcal{P}}_n$, are positively represented, i.e. $\xi_a^\gamma(\Pi_{s,m}) \geq 0$, $\forall a$ and $\forall \gamma \in \gamma_n$.*

Proof. The Wigner Function representation for measurement projectors is given by;

$$\xi_a^\gamma(\Pi_{s,m}) = \mathrm{Tr}\left(B_a^\gamma \Pi_{s,m}\right).$$

The set of phase-point operators, $\{B_a^\phi\}$ can be expressed as;

$$B_0^\gamma = \frac{1}{2^n} \sum_b (-1)^{\gamma(b)} P_b,$$

$$B_a^\gamma = P_a B_0^\gamma P_a^\dagger = \frac{1}{2^n} \sum_b (-1)^{\gamma(b)+[a,b]} P_b,$$

where $P_a \in \widetilde{\mathcal{P}}_n$. Using these two definitions we have;

$$\xi_a^\gamma(\Pi_{s,m}) = \text{Tr}\left(B_a^\gamma \Pi_{s,m}\right),$$

$$= \text{Tr}\left(\frac{1}{2 \cdot 2^n} \sum_b (-1)^{\gamma(b)+[a,b]}\left(\mathbb{I} + (-1)^s P_m\right) P_b\right),$$

$$= \frac{1}{2}\left(1 + \sum_b (-1)^{s+\gamma(b)+[a,b]}\delta_{m,b},\right)$$

$$= \frac{1}{2}\left(1 + (-1)^{s+\gamma(m)+[a,m]}\right) = \delta_{s,\,\gamma(m)+[a,m]} \geq 0, \; \forall a, \text{ and } \forall \gamma \in \gamma_n.$$

∎

This theorem tells us that every Wigner function W^γ represents all Pauli measurements positively and outcome-deterministically. This is actually unsurprising given the connection between quasi-probability representations and non-contextual ontological models. However we shall now show that the frame-switching model breaks down when we try to model measurement-update, after a measurement of a Pauli observable. To do so we will need the following lemma;

Lemma 4.4.7 *Given a measurement of observable $P_m \in \widetilde{\mathcal{P}}_n$, with outcome s, the phase-point operator A_a^γ is projected onto;*

$$\Pi_{s,m} A_a^\gamma \Pi_{s,m} = \frac{1}{4}\left(A_a^\gamma + A_{a+m}^\gamma + A_a^{\gamma'} + A_{a+m}^{\gamma'}\right), \tag{4.40}$$

where $\gamma'(b) = \gamma(b+m) + \gamma(m) + \beta(b,m), \forall b | [b,m] = 0$ and $\gamma'(b) = \gamma(b)$ otherwise. The function $\beta(b,m)$ is defined via $P_b P_m = (-1)^{\beta(b,m)} P_{b+m} | [b,m] = 0$.

Proof. From the definition of the phase-point operators and using $\Pi_{s,m} = \frac{1}{2}(\mathbb{I} + (-1)^s P_m)$,

we have;

$$\Pi_{s,m} A_a^\gamma \Pi_{s,m} = \frac{\mathbb{I} + (-1)^s P_m}{2} A_a^\gamma \frac{\mathbb{I} + (-1)^s P_m}{2},$$

$$= \frac{1}{4 \cdot 4^n} \sum_b (-1)^{\gamma(b)+[a,b]} \left(P_b + (-1)^s (P_m P_b + P_b P_m) + P_m P_b P_m \right),$$

$$= \frac{1}{\cdot 4^n} \sum_{b| [b,m]=0} \left(\frac{\mathbb{I} + (-1)^s P_m}{2} \right) (-1)^{\gamma(b)+[a,b]} P_b,$$

$$= \frac{1}{2 \cdot 4^n} \sum_{b| [b,m]=0} (-1)^{\gamma(b)+[a,b]} P_b + (-1)^{\gamma(b)+[a,b]+s+\beta(b,m)} P_{b+m}.$$

Splitting the above expression into the first and second terms of the sum, we have or the first term we have;

$$\sum_{b| [b,m]=0} (-1)^{\gamma(b)+[a,b]} P_b = \frac{4^n}{2} \left(A_a^\gamma + A_{a+m}^\gamma \right).$$

For the second term we use theorem 4.4.6 to note that given the point a in phase space the outcome is given by $s = \gamma(m) + [a,m]$, and by the nature of update maps Γ_Π is only defined for points such that $\xi_\Pi > 0$, to avoid counterfactual updates. Therefore the second term becomes;

$$\sum_{b| [b,m]=0} (-1)^{\gamma(b)+[a,b]+s+\beta(b,m)} P_{b+m} = \sum_{b| [b,m]=0} (-1)^{\gamma(b)+\gamma(m)+[a,b+m]+\beta(b,m)} P_{b+m},$$

$$= \sum_{b| [b,m]=0} (-1)^{\gamma(b+m)+\gamma(m)+\beta(b,m)+[a,b]} P_b,$$

$$= \frac{4^n}{2} \left(A_a^{\gamma'} + A_{a+m}^{\gamma'} \right),$$

where $\gamma'(b) = \gamma(b+m) + \gamma(m) + \beta(b,m), | [b,m] = 0$ and $\gamma'(b) = \gamma(b)$ otherwise. To move from line 1 to 2 we have also applied the identity $\beta(b+m,m) = \beta(b,m)$;

$$(-1)^{\beta(b+m,m)} P_b = (-1)^{\beta(b+m,m)} P_{m+b+m} = P_m P_{b+m} = (-1)^{\beta(b,m)} P_m P_b P_m$$

$$= (-1)^{\beta(b,m)+[b,m]} P_b = (-1)^{\beta(b,m)} P_b$$

Putting the above together we have;

$$\Pi_{s,m} A_a^\gamma \Pi_{s,m} = \frac{1}{4} \left(A_a^\gamma + A_{a+m}^\gamma + A_a^{\gamma'} + A_{a+m}^{\gamma'} \right), \tag{4.41}$$

■

So what does lemma 4.4.7 tell us? Well some of the literature out there would suggest that this lemma implies that under a sampling regime (i.e. a is a sample from a positive distribution) that we randomly update our frame from γ to γ or γ' (equally $a \mapsto \{a, a + m\}$). However this interpretation fails to give positive update maps;

> **Theorem 4.4.8** *Given some Wigner function W^γ there exists a $P \in \widetilde{\mathcal{P}}_n$ such that $\Gamma_{(P,s),\gamma \to \gamma} < 0$ and $\Gamma_{(P,s),\gamma \to \gamma'} < 0$.*

Proof. We will prove this statement by example, with the generalization coming from the application of a Clifford to the circuit to map to other phase-functions. We choose the 2-qubit Wigner function W^{γ_0}[16], i.e the Wigner function where $\gamma_0(b) = 0$, $\forall b$.

Under this choice of Wigner function the state $|00\rangle$ is positively represented. If we perform a XX measurement on this state and post-selecting on the $+1$ outcome, the post-measurement state will be $|B_{00}\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$. However this state is negatively represented in both choices of frame given by lemma 4.4.7, where $\gamma_0'(b) = \beta(b, XX)$, $\forall b | [b, XX] = 0$ and $\gamma_0'(b) = 0$ otherwise, $W^{\gamma_0}(|B_{00}\rangle \langle B_{00}|) < 0$ and $W^{\gamma_0'}(|B_{00}\rangle \langle B_{00}|) < 0$. This implies that the update maps must have negative elements, i.e. $\Gamma_{(XX,+1),\gamma_0 \to \gamma_0} < 0$ and $\Gamma_{(XX,+1),\gamma_0 \to \gamma_0'} < 0$.

This proof can then be generalized to any other frame via applying a Clifford to the above circuit, using;

$$\Pi_{XX,+1} |00\rangle \langle 00| \Pi_{XX,+1} \propto |B_{00}\rangle \langle B_{00}|$$
$$\longrightarrow C\Pi_{XX,+1}C^\dagger C |00\rangle \langle 00| C^\dagger C\Pi_{XX,+1}C^\dagger \propto C |B_{00}\rangle \langle B_{00}| C^\dagger$$

I.e. we map $\gamma_0 \to \gamma_C$ via the application of some Clifford C. This implies, by theorem 4.4.5, that;

$$W^{\gamma_C}(C |00\rangle \langle 00| C^\dagger) \geq 0,$$
$$W^{\gamma_C}(C |B_{00}\rangle \langle B_{00}| C^\dagger) < 0,$$
$$W^{\gamma_C'}(C |B_{00}\rangle \langle B_{00}| C^\dagger) < 0.$$

Therefore the measurement update maps $\Gamma_{(C(XX)C^\dagger,\gamma_C(XX)),\gamma_C \to \gamma_C}$ and $\Gamma_{(C(XX)C^\dagger,\gamma_C(XX)),\gamma_C \to \gamma_C'}$ are negatively represented, where we have changed the outcome to $\gamma_C(XX)$. ∎

---

[16]This proof can easily be embedded for larger n.

With this necessity of negatively represented measurement update maps we lose all hope of a positively represented model (and furthermore efficient simulation scheme). The attempts to rectify this and the conclusions it lead us to are covered in the final section of this chapter. Finally, we note that the update maps $\Gamma_{(P,s),\gamma\to\gamma}$ and $\Gamma_{(P,s),\gamma\to\gamma'}$ are equal. This is because eq. (4.40) is symmetric on γ and γ' and therefore via an easy calculation $\Gamma_{(P,s),\gamma\to\gamma} = \Gamma_{(P,s),\gamma\to\gamma'}$. This implies that switching the frame actually has no effect on the representation of the post-measurement state.

4.5 The β-function in the Binary Symplectic Representation

While the above theorems and lemmas provide a rigorous method for studying the frame-switching Wigner function. It is often the case that numerics are needed to verify the proofs, and provide graphical representations of the Wigner function. To do so, we use binary-symplectic representation to reduce the computational complexity of the Wigner function, considering that each Wigner function is defined on a 4^n grid. Crucially we need a closed form expression for the β-function in the binary-symplectic representation. Recalling that the β-function is defined as;

$$P_a P_b = (-1)^{\beta(a,b)} P_{a+b} | [a,b] = 0, \forall P_a, P_b, P_{a+b} \in \widetilde{\mathcal{P}}_n. \tag{4.42}$$

However, to derive an expression for the function it will be useful to drop the commuting assumption and use;

$$P_a P_b = i^{\phi(a,b)} P_{a+b}.$$

Recalling eq. (4.19) we can express any projective Pauli operator as;

$$P_a = i^{a_x \cdot a_z} X(a_x) Z(a_z) = \bigotimes_i i^{a_{x_i} a_{z_i}} X^{a_{x_i}} Z^{a_{z_i}}. \tag{4.43}$$

Noting that the dot product $a_x.a_z$ effectively is a counter for the number of Y factors in the tensor product. From this we can directly derive an expression for this ϕ function,;

$$P_a P_b = i^{a_x.a_z+b_x.b_z} \bigotimes_i^n X^{a_{x_i}} Z^{a_{z_i}} X^{b_{x_i}} Z^{b_{z_i}},$$

$$= i^{a_x.a_z+b_x.b_z} (-1)^{a_z.b_x} \bigotimes X^{a_{x_i}+b_{x_i}} Z^{a_{z_i}+b_{z_i}},$$

$$= i^{a_x.a_z+b_x.b_z+2a_z.b_x-(a_x+b_x).(a_z+b_z)} \bigotimes i^{(a_{x_i}+b_{x_i}).(a_{z_i}+b_{z_i})} X^{a_{x_i}+b_{x_i}} Z^{a_{z_i}+b_{z_i}},$$

$$= i^{a_z.b_x-a_x.b_z} \bigotimes i^{(a_{x_i}+b_{x_i}).(a_{z_i}+b_{z_i})} X^{a_{x_i}+b_{x_i}} Z^{a_{z_i}+b_{z_i}}$$

Now if we re-investigate eq. (4.43) we note that by assumption $a_{x_i}, a_{z_i} \in \mathbb{Z}_2$ and therefore $(a_{x_i} a_{z_i}) \in \mathbb{Z}_2$. However, in the above we have $(a_{x_i} + b_{x_i}).(a_{z_i} + b_{z_i}) \in \{0, 1, 2\}$, to convert it to the correct form we use;

$$a + b \bmod 2 = a + b - 2ab, \ a, b \in \mathbb{Z}_2.$$

Focusing on the i exponent in the tensor product we have;

$$\begin{aligned}
(a_{x_i} + b_{x_i}).(a_{z_i} + b_{z_i}) &= ([a_{x_i} + b_{x_i}] \bmod 2 + 2a_{x_i}b_{x_i}).([a_{z_i} + b_{z_i}] \bmod 2 + 2a_{z_i}b_{z_i}), \\
&= ([a_{x_i} + b_{x_i}] \bmod 2).([a_{z_i} + b_{z_i}] \bmod 2) \\
&\quad + 2a_{x_i}b_{x_i}[a_{z_i} + b_{z_i}] \bmod 2 + 2a_{z_i}b_{z_i}[a_{x_i} + b_{x_i}] \bmod 2, \quad (4.44) \\
&= ([a_{x_i} + b_{x_i}] \bmod 2).([a_{z_i} + b_{z_i}] \bmod 2) \\
&\quad + 2a_{x_i}b_{x_i}(a_{z_i} + b_{z_i} - 2a_{z_i}b_{z_i}) + 2a_{z_i}b_{z_i}(a_{x_i} + b_{x_i} - 2a_{x_i}b_{x_i}), \\
&= ([a_{x_i} + b_{x_i}] \bmod 2).([a_{z_i} + b_{z_i}] \bmod 2) \\
&\quad + 2a_{x_i}b_{x_i}(a_{z_i} + b_{z_i}) + 2a_{z_i}b_{z_i}(a_{x_i} + b_{x_i}). \quad (4.45)
\end{aligned}$$

Where we have repeatedly used the fact that anything with a factor of 4 disappears.

Finally re-substituting the above into the previous derivation, noting that $a_z.b_x - a_x.b_z = [a, b]$ which is the binary symplectic inner product, and defining a triple product of vectors as $a.b.c = \sum_i a_i b_i c_i$, we get;

$$P_a P_b = i^{[a,b]+2a_x.b_x.(a_z+b_z)+2a_z.b_z.(a_x+b_x)} P_c, \quad (4.46)$$

and therefore;

$$\phi(a, b) = [a, b] + 2a_x.b_x.(a_z + b_z) + 2a_z.b_z.(a_x + b_x), \quad (4.47)$$

$$\beta(a, b) = \frac{1}{2}[a, b] + a_x.b_x.(a_z + b_z) + a_z.b_z.(a_x + b_x) \,|\, [a, b] \in 0, 2, \quad (4.48)$$

where we note the binary symplectic product indicates two Pauli operators commute if $[a, b] = \{0, 2\}$.

Finally, if we define $P_a P_b P_c = (-1)^{\beta(a,b,c)} \mathbb{I} \mid [a, b] = 0$, where $c_x = a_x + b_x \bmod 2$ and $c_z = a_z + b_z \bmod 2$, the β function can be expressed as, via eq. (4.44), as;

$$\beta(a, b, c) = \frac{1}{2}[a, b] + a_x.b_x.c_z + a_z.b_z.c_x,$$

which could have a useful operational interpretation.

4.6 Relation to the Phase-Space Simulation by Raussendorf et Al.

Recently, Raussendorf et al. presented a quasi-probability simulation of the n-qubit stabilizer formalism [69] that shares many of the features of the frame-switching Wigner function and the ψ-epistemic model presented in chapter 5. In this section I shall provide a connection between the three approaches.

First let us briefly overview the Raussendorf et al construction, where we will restrict ourselves to qubits. Firstly in their construction they define phase-point operators of the form, under the notation given in this chapter;

$$A_\Omega^\gamma = \frac{1}{2^n} \sum_{b \in \Omega} (-1)^{\gamma(b)} P_b, \tag{4.49}$$

where $\Omega \subset \widetilde{\mathcal{P}}_n$. Hence the ontology of their quasi-probability representation is given by the pair (Ω, γ). Both Ω and γ are then defined to satisfy certain properties. Most importantly, Ω is *closed under inference*; if P_a and P_b are in Ω and commute then P_{a+b} is also in Ω. And γ satisfies a *non-contextual value assignment* for all a and b in Ω; i.e. γ satisfies $\gamma(a) + \gamma(b) + \gamma(a + b) = \beta(a, b)$. Lastly, they show that the set of Ω and γ that satisfy these properties can be parameterized by an integer m such that Ω has a center[17] of size 2^{n-m}.

While the Raussendorf et al paper does not give an explicit proof of these facts I believe the frame they define is self-dual, by using a modified version of lemma 4.4.1, and

---

[17]Where the center is defined as the largest subset of Ω such that all elements of Ω commute with the center.

all stabilizer states are positively represented, via a modified theorem 4.4.3. However, it should be noted that in their paper they leave the quasi-probability formalism behind by letting density operators be decomposed into arbitrary sets of phase-point operators, such that a positive decomposition can be found. At a glance this procedure would appear to reproduce the quantum statistics. However, I believe it would be prudent to verify this method as it could be the case that the frame's normalization and definition of the dual frame might cause problems. Additionally, by considering decompositions of this form their model's mapping from quantum states to quasi-probability distributions is no longer linear.

For the following discussion we will treat the Raussendorf et al. construction as a linear QPR[18]. Firstly we note that the sets Ω play a very similar role to the different frames in the Wigner function presented above. Namely that if we fix an Ω then we can define an origin and use Pauli translation operators to move us between the different γ associated with Ω. I.e. the set $\Gamma(\Omega)$ plays the same role as the equivalence classes $\mathcal{C}_\gamma$. Where the definitions differ though is the requirement of a non-contextual assignment, with the frame-switching Wigner function above imposing no restrictions on the structure of γ.

The ψ-epistemic model presented in chapter 5 was derived by modifying the frame-switching Wigner function presented in this chapter. As a consequence it retains many of the features of the Wigner function. The version of the ψ-epistemic model we shall use here is the reduction given in section 5.9. This reduction reduces the size of the ontology and converts the model from an outcome deterministic model to a outcome in-deterministic model, much like the Raussendorf et al. simulation. This reduction's ontology is given by the tableau;

$$
\lambda \equiv
\begin{array}{c|cccc|cccc|c}
 & a_{x_1,g_1} & a_{x_2,g_1} & \cdots & a_{x_n,g_1} & a_{z_1,g_1} & a_{z_2,g_1} & \cdots & a_{z_n,g_1} & \gamma(g_1) \\
 & a_{x_1,g_2} & a_{x_2,g_2} & \cdots & a_{x_n,g_2} & a_{z_1,g_2} & a_{z_2,g_2} & \cdots & a_{z_n,g_2} & \gamma(g_2) \\
\{g_i\}_i & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
 & a_{x_1,g_{n-1}} & a_{x_2,g_{n-1}} & \cdots & a_{x_n,g_{n-1}} & a_{z_1,g_{n-1}} & a_{z_2,g_{n-1}} & \cdots & a_{z_n,g_{n-1}} & \gamma(g_{n-1}) \\
\hline
 & a_{x_1,S_1} & a_{x_2,S_1} & \cdots & a_{x_n,S_1} & a_{z_1,S_1} & a_{z_2,S_1} & \cdots & a_{z_n,S_1} & \gamma(S_1) \\
\{S_i\}_i & a_{x_1,S_2} & a_{x_2,S_2} & \cdots & a_{x_n,S_2} & a_{z_1,S_2} & a_{z_2,S_2} & \cdots & a_{z_n,S_2} & \gamma(S_2) \\
 & a_{x_1,S_3} & a_{x_2,S_3} & \cdots & a_{x_n,S_3} & a_{z_1,S_3} & a_{z_2,S_3} & \cdots & a_{z_n,S_3} & \gamma(S_3)
\end{array}
\tag{4.50}
$$

Where $\{g_i\}_i$ represent a stabilizer group of size $n-1$ and $\{S_i\}_i$ are the three possible commuting extensions to the stabilizer group. The first thing we note is that this tableau is "non-contextual" and "closed under inference" as defined in [69], if we consider it to

---

[18]With the connections still holding when this assumption is dropped.

define a set $\Omega = \langle g_1, g_2, \ldots, g_{n-1}, S_1, S_2, S_3 \rangle$ and a phase-function γ equal to the phases in the tableau.

Considering the above I conjecture that the reduced ψ-epistemic model is equivalent to the Raussendorf QPR[19] which is constructed from all phase-point operators given in eq. (4.49) such that $m = 1$.
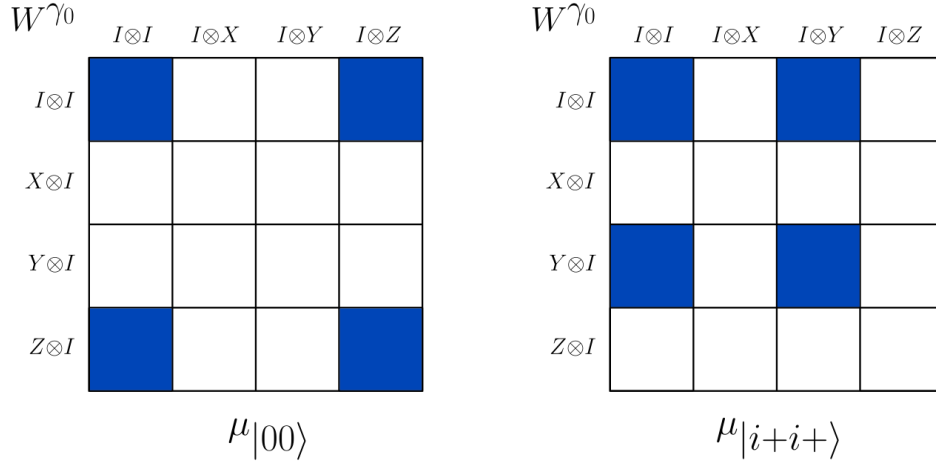
4.7 Final Remarks

From these results we can see that some kind of unified frame-switching Wigner function, which would play an analogous role to the n-qudit Wigner function, is dead in the water. However, there are still some questions that may be asked. For example, theorem 4.4.8 does not say that *all* Pauli measurement updates are negatively represented. This leads us to the question; maybe we can limit the Pauli's we can measure, given a frame, and use Clifford operations to reach the rest. However as shall be seen by the results in Chapter 6 this is not possible. We can always construct some sequence of measurements that will eventually require a negatively represented map.

So, what can we salvage from this project? Firstly let us investigate how a weak-simulation of the n-qubit stabilizer subtheory would be performed using the frame-switching Wigner function. Firstly we would prepare some stabilizer state, ψ, in a Wigner function, W^γ, such that $W^\gamma(\rho_\psi) \geq 0$. As the distribution is uniform we can then sample a point in phase-space from this distribution to perform our weak simulation with, let us call this point A_a^γ such that $W_a^\gamma(\rho_\psi) > 0$. Then by using lemma 4.4.4 and theorem 4.4.5 we can evolve our sample to a new point in phase-space and update our phase function accordingly, i.e. $(a, \gamma) \mapsto (C(a), C(\gamma))$. When we reach a measurement in the circuit we can use theorem 4.4.6 to output an outcome. Finally, the measurement update is trying to move us to a frame where the post-measurement state is positively represented. However, there are ambiguities in how this update should be performed, due to the amount of information contained in our sample a.

To understand how this ambiguity arises it is useful to return to the example used in theorem 4.4.8. Figures 4.1a and 4.1b gives the representations of the $|00\rangle$ and $|i + i+\rangle$ state in the γ_0 Wigner function. As can be seen both states share support on the origin, $A_0^{\gamma_0}$. Additionally, the origin also gives a $+1$ outcome for a measurement of XX, as shown in fig. 4.2, satisfying the example in theorem 4.4.8. Looking at how the stabilizer groups

---

[19]Under the linear definition of a QPR.

W^{γ_0}	$I \otimes I$	$I \otimes X$	$I \otimes Y$	$I \otimes Z$
$I \otimes I$	■			■
$X \otimes I$				
$Y \otimes I$				
$Z \otimes I$	■			■

$\mu_{|00\rangle}$

W^{γ_0}	$I \otimes I$	$I \otimes X$	$I \otimes Y$	$I \otimes Z$
$I \otimes I$	■		■	
$X \otimes I$				
$Y \otimes I$	■		■	
$Z \otimes I$				

$\mu_{|i+i+\rangle}$

(a) The Wigner function representation of $|00\rangle$, where we have chosen $\gamma(b) = 0, \forall b$.

(b) The Wigner function representation of $|i + i+\rangle$, where we have chosen $\gamma(b) = 0, \forall b$.

Figure 4.1: A graphical representation of the W^{γ_0} Wigner function. The blue boxes indicate phase-points with, positive, non-zero support for each given state. The Pauli operators represent the translation operators required to reach a phase-point from the origin, located top left.

of $|00\rangle$ and $|i + i+\rangle$ update after the measurement of XX we have;

$$|00\rangle : \{\mathbb{I}, \mathbb{I}Z, Z\mathbb{I}, ZZ\} \xrightarrow{XX,+1} \{\mathbb{I}, ZZ, XX, -YY\},$$

$$|i + i+\rangle : \{\mathbb{I}, \mathbb{I}Y, Y\mathbb{I}, YY\} \xrightarrow{XX,+1} \{\mathbb{I}, YY, XX, -ZZ\}.$$

Noting theorem 4.3.1 we can see that after the measurement of XX, with a +1 outcome $|00\rangle$ and $|i + i+\rangle$ are mapped to orthogonal states. This implies, by another of our results [72], that the representations of $|00\rangle$ and $|i + i+\rangle$ must share no support over the support of $\xi_{\Pi_{XX},+1}$. This is clearly not the case in this Wigner function, and therefore to avoid a contradiction negativity must be introduced after the measurement.

However digging a bit deeper we can start to see what the Wigner function is trying to achieve during measurement update. From theorem 4.4.3 we can see that a stabilizer state is positively represented *iff* the phase-function is equal to the phases on the stabilizer elements, up to some balanced vector. Implying the Bell basis, $|B_{00}\rangle$ etc., is negatively represented in the all-zero phase function. However it will be positively represented in any
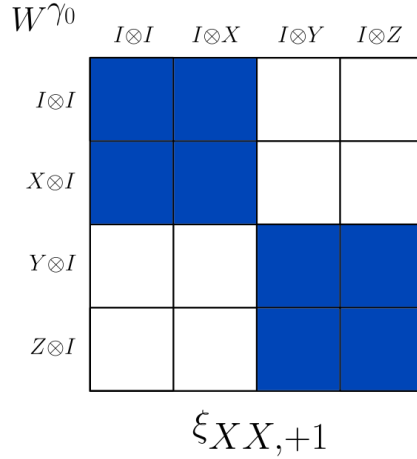
$$W^{\gamma_0}$$

$$\xi_{XX,+1}$$

Figure 4.2: A graphical representation of the W^{γ_0} Wigner function's representation of the projector $\Pi_{XX,+1} = (\mathbb{I} + XX)/2$. The blue boxes indicate phase-points with, positive, non-zero support. The Pauli operators represent the translation operators required to reach a phase-point from the origin, located top left.

phase function that satisfies;

$$\gamma(XX) + \gamma(ZZ) + \gamma(YY) \bmod 2 = 1 = \beta(XX, ZZ). \tag{4.51}$$

Noting that any balanced vector over the Pauli group defined as $v(a) = [a, c]$ for some $c \in \mathcal{B}(\widetilde{\mathcal{P}}_n)$, is also balanced over any Abelian subgroup of $\widetilde{\mathcal{P}}_n$.

Therefore after a measurement of XX the Wigner function tries to move to a phase-function that satisfies the above relationship. However it runs into a problem, how to set the phase of YY and ZZ? Without extra knowledge of the state of the system prior to measurement (say by knowing an element of the stabilizer group, not equal to the identity), it tries to set the phases for YY and ZZ for both possibilities. Setting $\gamma_0'(YY) = \beta(XX, YY) = 1$ to satisfy the update $|00\rangle \xrightarrow{XX,+1} |B_{00}\rangle$ and setting $\gamma_0'(ZZ) = \beta(XX, ZZ) = 1$ to satisfy $|i + i+\rangle \xrightarrow{XX,+1} |B_{10}\rangle$. Which when substituted into eq. (4.51) fails to generate a phase-function that positively represents the post-measurement state, where $\gamma_0'(XX) = \gamma_0(XX)$.

Armed with this knowledge we can start to build a ψ-epistemic model of the stabilizer formalism, which is explicitly constructed to avoid ambiguities of the form presented above. This is the model presented in chapter 5. Considering these ambiguities in general will then lead us to the results of chapter 6, which tell us that *any* model of the n-qubit stabilizer

formalism requires at least $n-1$ generators to be encoded in the ontology to correctly reproduce state update. This represents one of the most stringent requirements we could place on a ψ-epistemic theory, and strongly suggests that any model of full quantum theory must be ψ-ontic.

Chapter 5

A Contextual ψ-Epistemic Model of the n-Qubit Stabilizer Formalism

5.1 Chapter Preamble

In this chapter I present work that followed on from the previous chapter's work on frame switching Wigner functions. The frame switching methodology was unable to represent measurement update correctly and to make any progress I had to approach the project from a different angle. To do so I began by stripping away the Wigner function machinery and started to investigate a model that directly sampled a stabilizer operator from the stabilizer group of a state (effectively what the sampling part of the frame switching Wigner function was doing) and deterministically stored outcomes of measurements (as the phase function in the Wigner function does).

   This approach lead to the first iteration of this model. However, it soon became clear that the single-stabilizer sampling could not reproduce the quantum statistics. To avoid this problem it became necessary to encode $n-1$ generators in the ontology, such that measurement update could be performed consistently. This would then be extended to show that *any* ontological model of the n-qubit stabilizer formalism must encode $n-1$ generators, which is presented in chapter 6.

   The work presented in this paper was completed solely by myself, with supervision from Joseph Emerson. It is currently published on the arXiv, with submission to journals to follow [52].

A Contextual ψ-Epistemic Model of the n-Qubit Stabilizer Formalism

Piers Lillystone[1], and Joseph Emerson[2,3]

[1]*Institute for Quantum Computing and Department of Physics and Astronomy, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada*
[2]*Institute for Quantum Computing and Department of Applied Mathematics, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada*
[3] *Canadian Institute for Advanced Research, Toronto, Ontario M5G 1Z8, Canada*

Contextuality, a generalization of non-locality, has been proposed as *the* resource that provides the computational speed-up for quantum computation. For universal quantum computation using qudits, of odd-prime dimension, contextuality has been shown to be a necessary and possibly sufficient resource [39]. However the role of contextuality in quantum computation with qubits remains open. The n-qubit stabilizer formalism, which by itself cannot provide a quantum computer super-polynomial computational advantage over a classical counterpart [31], is contextual. Therefore contextuality cannot be identified as a sufficient resource for quantum computation. However it can be identified as a necessary resource [71]. In this paper we construct a contextual ψ-epistemic ontological model of the n-qubit stabilizer formalism, to investigate the contextuality present in the formalism. We demonstrate it is possible for such a model to be outcome deterministic, and therefore have a value assignment to all Pauli observables. In our model the value assignment must be updated after a measurement of any stabilizer observable. Remarkably, this includes the value assignments for observables that commute with the measurement. A stronger manifestation of contextuality than is required by known no-go theorems.

5.2   Introduction

Contextuality is well-established as a fundamental non-classical feature of quantum theory [7, 43, 55, 63, 56, 75], including as a special case Bell non-locality. This has naturally lead to it being proposed as a candidate for understanding the source of quantum computational speed-up. Recent work in classifying contextuality as a resource for quantum computation has proved a fruitful research field. Howard *et al.* [39] proved that contextuality is a

necessary and possibly sufficient resource[1] for universal quantum computation, in the context of magic-state injection schemes, with odd-prime dimension qudits. For qubit quantum computation contextuality has been proven to be a necessary resource [25, 9, 71]. Further it can even quantify the amount of computational advantage in magic-state injection schemes and measurement-based models of quantum computation [4, 70, 82, 68, 83, 9, 61, 3, 21].

However, despite recent progress in identifying contextuality as a resource for quantum computation, it can only ever be considered as a necessary resource. This is because the n-qubit stabilizer formalism, a fundamental subset (which we call a subtheory) of quantum operations that is not universal for quantum computation [30, 15, 32], admits an efficient classical simulation [31, 81, 2, 62], but is contextual under any standard definition of contextuality [55, 63, 56, 4, 38, 45, 51]. Therefore contextuality cannot be classified as sufficient for universal quantum computation.

This motivates the question: can the concept of contextuality be re-characterized or refined such that a revised notion can be identified as both a necessary and sufficient resource for universal quantum computation? If such a redefinition is possible it would unify the foundational principle that contextuality is the fundamental non-classical property of quantum theory [7, 43, 56] which justifies the widely held belief that quantum computation is more powerful than classical computation [64]. A natural next step toward answering this question is to understand exactly how contextuality manifests in the n-qubit stabilizer formalism by constructing an explicit contextual ontological model of that subtheory.

Within the landscape of ontological models there are two primary types [35, 36, 50]: ψ-ontic ontological models wherein a quantum state is uniquely specified by the physical state, and ψ-epistemic ontological models wherein quantum states are not fully specified by the underlying physical state of system, i.e. there exist physical states that can be prepared by multiple quantum states. A ψ-epistemic ontological model provides for a classical-like model of quantum theory and is appealing because it provides a means to evade the measurement problem while maintaining a classically natural notion of objective reality. For example, the Kochen-Spekker model of a qubit [43], the 8-state model of the qubit stabilizer formalism [84, 11], and Gaussian quantum mechanics [6] are all ψ-epistemic models of subsets of quantum theory. Conversely, in a ψ-ontic model the quantum state fully describes the state of a system and hence cannot give a classically natural account of quantum theory. It remains an open question whether a physically plausible ψ-epistemic model of quantum theory is possible, with recent work [72] demonstrating that all previously known ψ-epistemic models of quantum theory cannot reproduce state

---

[1]With sufficiency dependent on the conjecture [83] that all $\rho \notin \mathcal{P}_{\mathrm{SIM}}$, where $\mathcal{P}_{\mathrm{SIM}}$ is the set of ancilla states not useful for magic state distillation routines, can be distilled to a magic state.

update.

The Gottesman-Knill theorem can be used to define a ψ-ontic ontological model of the n-qubit stabilizer subtheory. This is also the case for the computational basis phase model [24, 81], which can be used to construct a weak-simulation of the n-qubit stabilizer subtheory and beyond [14]. For the, non-contextual, n-qudit stabilizer subtheory, of odd prime dimension, a ψ-epistemic model is possible [79, 21]. However, for qubits the only previously known ψ-epistemic model was for the single qubit stabilizer formalism, the 8-state model [84]. Indeed, for a single qubit, all of quantum theory is traditionally non-contextual [43].

Hence it has been an interesting open question whether an ψ-epistemic ontological model of the full n-qubit stabilizer subtheory can be constructed, which is expected because we know that this subtheory admits an efficient classical description. In this paper we answer this open question in the positive by explicitly constructing a contextual ψ-epistemic model of the n-qubit stabilizer subtheory. We expect that this model will serve as a fundamental insight into the structure of contextuality within quantum theory and help clarify the role of contextuality as a resource for quantum speed-up for which the stabilizer subtheory serves as a stepping stone.

As we will see, the ψ-epistemic ontological model constructed in this paper is outcome deterministic, that is, physical states of the system fully encode the outcome of any stabilizer measurement. Outcome determinism is often seen as a core motivating principle behind the traditional notion of contextuality, often referred to as Bell-Kochen-Specker contextuality [7, 43]. Therefore it of interest that the model presented here satisfies this principle, a feature shared by previous contextual explanations of the Mermin-Peres square [41, 47]. However the model does not satisfy the requirement that the possible static value assignments respect all functional relationships between sets of commuting observables. Rather it only requires that a given value assignment is consistent with the current state and uses measurement update rules to satisfy the functional relationship between outcomes of commuting observables. We also provide a restructuring of the model in the appendix, which has the appealing feature of being an *always-ψ-epistemic* model, but is no longer outcome deterministic.

In our model a correct update to the value assignments of the model requires the physical states to encode $n - 1$ generators of a stabilizer state. However, we highlight as an open question whether a generic ψ-epistemic model of the n-qubit stabilizer formalism requires "almost" complete knowledge of stabilizer group of a state.

5.3 Preliminaries

5.3.1 Ontological Models

To build a model of the stabilizer formalism we use the ontological models formalism. To define an ontological model we begin with the mathematical framework of operational theories. An operational theory provides a framework for predicting the observed outcomes of experimental procedures, which can be decomposed into a sequence of preparations, transformations, and measurements. Hence an operational theory gives a specification of the experimentally observed statistics $\Pr(k|P, T, M)$ [36].

An ontological model provides a description of a set of experimental statistics by supposing that there exists a well defined notion of a real physical state of a system, referred to as the ontic state λ, which encodes all physically knowable properties of the system. The ontological models formalism then describes preparations, transformations, and measurements in terms of probability densities, stochastic processes, and response functions over the set of ontic states, Λ, which we call the ontic space:

Preparations: A preparation P is represented in an ontological model by a probability density μ_P over the ontic space: $\mu_P : \Lambda \to [0, 1]$, where $\int_\Lambda \mu_P(\lambda)d\lambda = 1$.

Transformations: A transformation T is represented in an ontological model by a stochastic map Γ_T between ontic states: $\Gamma_T : \Lambda \times \Lambda \to [0, 1]$, where $\int_\Lambda \Gamma_T(\lambda', \lambda)d\lambda' = 1, \ \forall \lambda \in \Lambda$.

Measurements: A measurement of a POVM, M, is represented in an ontological model by a set of conditional probability distributions over outcomes of the POVM $\{\xi_M(k|\lambda)|\, k \in M\}$, where $\xi_M(k) : \Lambda \to [0, 1]$, and $\sum_k \xi_M(k|\lambda) = 1, \ \forall \lambda \in \Lambda$.

To reproduce the statistics of an operational theory $\Pr(k|P, T, M)$ the ontological model must satisfy:

$$\Pr(k|P, T, M) = \int_{\Lambda'} \int_\Lambda \xi_M(k|\lambda')\Gamma_T(\lambda', \lambda)\mu_P(\lambda)d\lambda d\lambda'. \tag{5.1}$$

For example in this paper we are interested in the operational theory defined by the n-qubit stabilizer formalism. I.e. convex mixtures of; n-qubit pure stabilizer state preparations, unitary transformations from the Clifford group, and projective measurements of

Pauli observables[2].

ψ-Epistemic models

The ontological models formalism provides a platform for investigating statements about how physical reality may be structured. A central question in any such investigation is whether a quantum state provides a unique description of physical reality, i.e. whether every ontic state λ can be associated to a single quantum state ψ. Ontological models where this is the case are termed ψ-ontic. Conversely, ontological models where there exists at least one ontic state that could be prepared by multiple quantum states, are termed ψ-epistemic:

> **Definition 5.3.1** *A ψ-**ontic** ontological model, $\mathcal{O}$, is one where all pure quantum states' probability distributions, $\mathcal{P}(\mathcal{O})$, are non-overlapping:*
>
> $$\int_\Lambda \mu_{P_\psi}(\lambda)\mu_{Q_\phi}(\lambda)d\lambda = 0, \; \forall P_\psi, Q_\phi \in \mathcal{P}(\mathcal{O}).$$
>
> *Otherwise it is ψ-**epistemic**.*

For example, in the ontological model defined by the Gottesman-Knill theorem every possible tableau of generators is identified as an ontic state. Hence each ontic state corresponds to exactly one stabilizer state, making the model ψ-ontic. Note, a stabilizer state can be represented by $2^{n(n-1)/2}\prod_{k=0}^{n-1}(2^{n-k}-1)$ tableaus [2], therefore there are as many ontic states associated to a single stabilizer state.

An alternative way to express the criteria for an ontological model to be ψ-epistemic is to notice that definition 5.3.1 requires that there exists at least one ontic state λ that is in the support of more than one quantum state's support:

> **Definition 5.3.2** *An ontological model is ψ-**epistemic** if there exists an ontic state λ that is in the support of more than one pure quantum state, i.e. mathematical we require:*
>
> $$\exists \lambda \in supp(\mu_{P_\psi}) \cap supp(\mu_{Q_\phi}), \; P_\psi, Q_\phi \in \mathcal{P}(\mathcal{O}),$$

---

[2]In the fragment notation the qubit stabilizer subtheory is defined as $\mathcal{F}_{stab,2^n} =< \mathcal{S}(\mathcal{D}(\mathcal{H}_{2^n})), Cl_{2^n}, \widetilde{\mathcal{P}}_n > $ [50].

> *where the support is defined as $supp(\mu_{P_\psi}) = \left\{ \lambda | \mu_{P_\psi}(\lambda) > 0, \lambda \in \Lambda \right\}.$*

Requiring that there only exists one ontic state in the support of more than one quantum state is a very weak condition. A more natural requirement of an ontological model is that all ontic states are an element of more than one quantum state's support[3], such a model is called an always-ψ-epistemic model;

> **Definition 5.3.3** *An ontological model is an **always-ψ-epistemic** if the model satisfies;*
>
> *$\forall \lambda \in \Lambda$ there exist two or more operationally distinct pure state preparations $P_\psi, Q_\phi \in \mathcal{P}(\mathcal{O})$ such that $\lambda \in supp(\mu_{P_\psi}) \cap supp(\mu_{Q_\phi})$.*

Additionally we can require that all ontic states are in the support of the same number of quantum states. We term ontological models that satisfy this **symmetric**-always-ψ-epistemic models.

One of the interesting features of always-ψ-epistemic models is their relation to weak simulations [81]. Ideally in a weak simulation the act of sampling should not give us enough information to calculate the final distribution we wish to sample an outcome from. Similarly in an always-ψ-epistemic model having knowledge of which ontic state was prepared is insufficient to determine the quantum state of the system and therefore the full experimental statistics. However both are guaranteed to agree with the experimental statistics over many repeated runs, and never disagree with the experimental certainties.

5.3.2   Traditional Contextuality

Traditional non-contextuality, originally proposed by Kochen-Specker and Bell [7, 43, 55, 63, 56], is defined by assuming it is possible to assign values to sets of commuting observables, such that functional relationships between observables are satisfied. So given an observable M and an ontic state λ, λ deterministically specifies the outcome of a measurement M. Mathematically we express this value assignment as $v_\lambda(M) \in \{E_k\}$, where $\{E_k\}$ is the set of eigenvalues of M indexed by k.

Within quantum theory if an observable M is simultaneously measurable with some set of mutually commuting observables $\{A, B, ..\}$ then any functional relationship between

---

[3]Such a requirement firmly places the role of the quantum state at the level of a state of knowledge, i.e. given an ontic state we cannot determine which quantum state that ontic state.

these observables, $f(M, A, B, ..) = 0$, will also be satisfied by the simultaneous eigen-values of these observables, so $f(E_k, E_a, E_b, ..) = 0$. The assumption of traditional non-contextuality then assumes that this functional relationship should also hold for any value assignment, i.e. $f(v_\lambda(M), v_\lambda(A), v_\lambda(B), ...) = 0, \forall \lambda \in \Lambda$.

The above assumption implicitly assumes that the value assignments should satisfy the functional relationship for *any* set of commuting observables, i.e. $\{M, A, B, ...\}$ and $\{M, L, M, ...\}$, even if such sets do not mutually commute. It is this assumption that the model presented here explicitly breaks from. We only require that the value assignment given by an ontic state is consistent with the quantum state prepared and not any set of commuting measurements. The functional relationships between commuting observables are then satisfied via a measurement update rule, which also ensures the post-measurement value assignment is consistent with the post-measurement state.

5.3.3 The n-Qubit Stabilizer Formalism

As we are interested in building a contextual ontological model of stabilizer circuits we restrict ourselves to the unitary/PVM n-qubit stabilizer subtheory. The operational theory we are interested in is therefore $p(k| \rho, C, M)$, where $\rho \in \mathcal{S}(\mathcal{H})$, $C \in Cl_{2^n}$, $M \in \widetilde{\mathcal{P}}_n$.

The most basic element of the stabilizer formalism is projective Pauli group $\widetilde{\mathcal{P}}_n$, which we define as;

> **Definition 5.3.4** *The **projective Pauli group** is the standard Pauli group, $\mathcal{P}_n$, modulo global phases;*
>
> $$\widetilde{\mathcal{P}}_n = \left\{ \bigotimes_i^n P_i \middle| P_i \in \{\mathbb{I}, X, Y, Z\} \right\} = \mathcal{P}_n/U(1).$$

We say an operator *stabilizes* a state if applying that operator to the state leaves it unchanged, i.e. it is an element of the $+1$ eigenspace of the operator. In the stabilizer formalism we choose the projective Pauli group with ± 1 phases as our set of *stabilizer operators*. Hence the set of stabilizer operators is isomorphic to $(\mathbb{Z}_2 \times \widetilde{\mathcal{P}}_n) \setminus -\mathbb{I} \cong \mathcal{S}_n$, where we have removed $-\mathbb{I}$ as it stabilizes no state. Note the full Pauli group, $\mathcal{P}_n$, contains operators that square to negative identity and therefore do not stabilize any state, i.e. non-Hermitian Pauli operators such as $iX \in \mathcal{P}_n$.

For the purposes of the model we are interested in pure stabilizer states. Mixed stabilizer states in the model are considered to be convex combinations of pure stabilizer states. The set of pure stabilizer states is defined as the set of pure states that are in the joint $+1$ eigenspace of n mutually commuting stabilizer operators, i.e. $\mathcal{S}(\psi) = \{S \,|\, S \,|\psi\rangle = |\psi\rangle,\, S \in \mathcal{S}_n\}$ where $[S, S'] = 0,\, \forall S, S' \in \mathcal{S}(\psi)$. From this it can be shown that the *stabilizer group* of $|\psi\rangle$, $\mathcal{S}(\psi)$, is a maximal abelian subgroup of the Pauli group [32], with n generators. Further the stabilizer group uniquely specifies the stabilizer state. To see this consider that if we define $\mathcal{S}(\psi) = \langle g_1, g_2, ..., g_n \rangle,\, g_i \in \mathcal{S}_n$ then;

$$\rho_\psi = \prod_{i=1}^{n} \frac{\mathbb{I} + g_i}{2} = \frac{1}{2^n} \sum_{S \in \mathcal{S}(\psi)} S, \tag{5.2}$$

i.e. we repeatedly project onto each generator's $+1$-eigen-subspace until we reach a 1-dimensional subspace.

The unitaries in the stabilizer formalism are unitaries from the Clifford group, Cl_{2^n}, which is defined as the normalizer of the Pauli group;

$$Cl_{2^n} = \left\{ U \in U(2^n) \,|\, U S U^\dagger \in \mathcal{S}_n,\, \forall S \in \mathcal{S}_n \right\}. \tag{5.3}$$

Therefore Clifford gates, such as the Hadamard, phase, and CNOT gates, strictly map stabilizer states to stabilizer states.

As previous stated the observables measurable in the n-qubit stabilizer subtheory are all projective Pauli operators, where other stabilizer operators can be measured via a classical post-processing bit flip. Therefore the operational statistics any ontological model of the n-qubit stabilizer formalism must reproduce can expressed as;

Lemma 5.3.1 *Given a stabilizer state $|\psi\rangle \in \mathcal{S}(\mathcal{H})$ and a Pauli observable $M \in \widetilde{\mathcal{P}}_n$ the expectation value of the observable is given by;*

$$\langle M \rangle_\psi = Tr(M \rho_\psi) = \begin{cases} 1: & M \in \mathcal{S}(\psi) \quad (C1), \\ 0: & \pm M \notin \mathcal{S}(\psi) \quad (C2), \\ -1: & -M \in \mathcal{S}(\psi) \quad (C3), \end{cases} \tag{5.4}$$

Where $\mathcal{S}(\psi)$ denotes the stabilizer group of $|\psi\rangle$.

Proof. Simply substituting in equation 5.2 to the LHS and using the trace-orthonormality of the projective Pauli operators, $\mathrm{Tr}(P_i P_j) = 2^n \delta_{i,j}$, retreives the desired result. ∎

Note here we have used the expectation value for clarity over the corresponding projectors. As the expectation is bounded between -1 and $+1$, and the observables $M \in \tilde{\mathcal{P}}_n$ have ± 1 eigenvalues, we can see that if $M \in \mathcal{S}(\rho)$ then the outcome must be 1 (C1), similarly -1 for $-M \in \mathcal{S}(\rho)$ (C3). Condition 2 (C2) implies the outcome of the measurement is uniformly random, over $\{+1, -1\}$, if $\pm M \notin \mathcal{S}(\rho)$.

5.3.4 The Stabilizer Formalism and Measurement Update Rules

Central to the model's construction are the ontological measurement update rules. Here we recap how measurement updates are performed in the stabilizer formalism.

As stabilizer measurements are rank 2^{n-1} PVMs the update rules for them are relatively straight forward to state. Suppose we measured the stabilizer state $|\psi\rangle$ with Pauli observable M and observed outcome $k \in \{0, 1\}$. Then the post measurement state is given by (with appropriate renormalization);

$$
\begin{aligned}
\rho_{\psi'} &\propto \frac{\mathbb{I} + (-1)^k M}{2} \rho_\psi \frac{\mathbb{I} + (-1)^k M}{2}, \\
&\propto \sum_{S \in \mathcal{S}(\psi)} S + (-1)^k (MS + SM) + MSM, \\
&\propto \sum_{S \in \mathcal{S}(\psi) | [S, M] = 0} S + (-1)^k MS, \\
&= \sum_{b \in \mathcal{B}(\mathcal{S}(\psi)) | [b, m] = 0} (-1)^{\gamma(b)} P_b + (-1)^{\gamma(b) + k + \beta(b, m)} P_{b+m},
\end{aligned} \tag{5.5}
$$

where we have used the definition $P_a P_b = (-1)^{\beta(a, b)} P_{a+b} | [P_a, P_b] = 0$. Expressing the update rule this way allows us to infer that the post-measurement stabilizer group is constructed by removing all elements of $\mathcal{S}(\psi)$ that do not commute with M and replacing them with compositions of M and the commuting group elements. This allows us to even further simplify the the update rule;

Given any stabilizer group of a pure state $\mathcal{S}(\psi)$ and any projective Pauli operator M, it is always possible to express the group as $\mathcal{S}(\psi) = \langle G, h \rangle$ where G is a proper subgroup of $\mathcal{S}(\psi)$ such that $[G_i, M] = 0, \forall G_i \in G$ and $\pm M \notin G$. Therefore the post-measurement group, after a measurement of M with outcome k, is homomorphic to $\mathcal{S}(\psi') = \langle G, (-1)^k M \rangle$, i.e. either $h \neq (-1)^k M$ and $[h, M] \neq 0$ and therefore we remove h and add $(-1)^k M$ in its place or $h = (-1)^k M$ and the pre-and-post measurement stabilizer groups are the same.

Many of known no-go theorems can be performed in the stabilizer formalism by using sequential measurement procedures. For example, the Mermin-Peres square, a proof of contextuality, and the PBR theorem, which can be implemented via adaptive stabilizer measurements.

5.4 A simple model via a global value assignment

Before building a full ontological model of the n-qubit stabilizer formalism, let us construct a model of a simpler subtheory. Namely the prepare-measure-discard n-qubit stabilizer subtheory, i.e. we allow preparations of pure stabilizer states and a terminating measurement of a Pauli observable. From the stabilizer statistics, lemma 5.3.1, the outcome of a measurement, M, only requires knowledge of whether $\pm M \in \mathcal{S}(\rho)$, which we can simply encode in a value assignment over all Pauli observables:

> **Definition 5.4.1** *The **value assignment** ν is a function that assigns $\{+1, -1\}$ to every element of the projective Pauli group $\widetilde{\mathcal{P}}_n$;*
>
> $$\nu : \widetilde{\mathcal{P}}_n \longrightarrow \{+1, -1\}, \tag{5.6}$$
>
> *such that $\nu(\mathbb{I}) = +1$.*

However for the majority of the this paper it will be more useful to use a binary representation of the value assignment, that is we map $+1 \rightarrow 0$ and $-1 \rightarrow 1$, which we term the *phase function*[4];

> **Definition 5.4.2** *The **phase function** γ is a function that assigns $\{0, 1\}$ to every element of projective Pauli group $\widetilde{\mathcal{P}}_n$;*
>
> $$\gamma : \widetilde{\mathcal{P}}_n \longrightarrow \mathbb{Z}_2, \tag{5.7}$$
>
> *such that $\gamma(\mathbb{I}) = 0$.*

To move between the two representations we can use $\nu(M) = (-1)^{\gamma(M)}$ and $\gamma(M) = \frac{1}{2}(1 - \nu(M))$. We denote the set of n-qubit phase functions as $\gamma_n = \left\{ \gamma \mid \gamma : \widetilde{\mathcal{P}}_n \longrightarrow \mathbb{Z}_2 \right\}$.

---

[4]Similar to the value assignments used in Wigner functions [25, 9, 71]

This definition naturally introduces the notion of consistency between stabilizer states and assignments γ;

> **Definition 5.4.3** *A phase function γ is **consistent** with a stabilizer state $\rho \in \mathcal{S}(\mathcal{H})$, denoted $\gamma \cong \rho$, with stabilizer group $\mathcal{S}(\rho) = \left\{ (-1)^{p_b} P_b | p_b \in \mathbb{Z}_2, \ P_b \in \widetilde{\mathcal{P}}_n \right\}$ iff;*
>
> $$\gamma(b) = p_b, \ \forall b \in \mathcal{B}(\mathcal{S}(\rho)),$$
>
> *where $\mathcal{B}$ denotes the binary sympletic representation of $\mathcal{S}(\rho)$.*

For the rest of this paper we will let lower case letters imply some index over the projective Pauli group such that they form a group homomorphism, for example the binary-sympletic representation.

Note in definition 5.4.3 the phase function retains consistency regardless of its value on points not in the stabilizer group. From this we can reproduce the stabilizer statistics by defining an epistemic state to be a uniform distribution over consistent phase functions and measurements to read-out the value of the phase function. Therefore the outcome of a measurement will be random if $\pm M \notin \mathcal{S}(\rho)$ and equal to the phase on M if $\pm M \in \mathcal{S}(\rho)$.

While the scope of the model is a highly restricted, what we can learn is that the phase function is sufficient to output the outcomes of measurements in the prepare and measure setting. The subsequently presented ψ-epistemic model of the n-qubit stabilizer formalism leverages this by effectively encoding the outcome of a measurement in the phase function. Then uses additional ontology to ensure that the post-measurement phase function is consistent with the post-measurement state, with randomization where necessary. Hence reproducing the quantum statistics.

5.5 A contextual ψ-Epistemic Model of the 2-Qubit Stabilizer Formalism

Before building the full n-qubit model it will be insightful to present the 2-qubit case, which has a slightly less abstract construction. Clearly to reproduce the outcome statistics of any prepare-transform-measure stabilizer circuit it is sufficient to only use the phase function as described earlier. However to handle sequential measurements the update rules requires additional ontology.

The additional ontology we require is non-trivial strict subgroups of the stabilizer group of the state. Which for two qubits is just the group containing one projective Pauli operator, $G = \left\{ \mathbb{I}, P | P \in \widetilde{\mathcal{P}}_2 \right\}$. We are only interested in projective Pauli operators as the phase function contains the information on whether P or $-P$ is an element of the stabilizer group.

This means the ontology of the 2-qubit model can be expressed as $\Lambda = \left(\widetilde{\mathcal{P}}_2 \backslash \mathbb{I} \right) \times \gamma_2$. The uniform distributions, μ_ψ, representing a pure stabilizer state $\psi \in \mathcal{S}(\mathcal{H}_{2^2})$ have support;

$$\operatorname{supp}(\mu_\psi) = \left\{ \lambda = (P, \gamma) | P \in \widetilde{\mathcal{S}}(\psi) \backslash \mathbb{I}, \gamma \cong \psi \right\}, \tag{5.8}$$

where the uniformity of the distribution ensures a random outcome for observables not in the stabilizer group. The model is ψ-epistemic as some non-orthogonal states have mutual support. For example, the computation basis state $|00\rangle$ has support on all $\lambda = (P, \gamma)$ such that $\gamma(\mathbb{I}Z) = \gamma(Z\mathbb{I}) = \gamma(ZZ) = 0$ and $P \in \{\mathbb{I}Z, Z\mathbb{I}, ZZ\}$. Therefore the all-zero state shares ontic states with a state $|\psi\rangle$ if $\exists P' \in \mathcal{S}(\psi)$ such that $P' \in \{\mathbb{I}Z, Z\mathbb{I}, ZZ\}$, and there exist phase functions consistent with both $|00\rangle$ and $|\psi\rangle$, i.e. $|0+\rangle$.

A Clifford operation's map on the ontology can be derived by considering it's map on the stabilizer operators, it's *stabilizer relations*. The stabilizer relations for a given Clifford operation C can be expressed as $C P_a C^\dagger = (-1)^{\gamma_c(a)} P_{c(a)}$ where $P_a, P_{c(a)} \in \widetilde{\mathcal{P}}_n$. Using this expression we represent a Clifford operation in the model by a permutation of the ontic states;

$$\Gamma_C : (P_a, \gamma) \longrightarrow (P_{c(a)}, \gamma^R + \gamma_c), \tag{5.9}$$

where γ_c is the vector of the phases $\gamma_c(a)$ and γ is reordered to match the Clifford permutation, therefore $(\gamma^R + \gamma_c)(a) = \gamma(c^{-1}(a)) + \gamma_c(a)$.

A measurement M's response functions "read-out" the value of the phase function;

$$\xi_{k,M}(\lambda) = \begin{cases} 1 & \text{if } \gamma(M) = k, \\ 0 & \text{otherwise.} \end{cases} \tag{5.10}$$

Therefore the model is outcome deterministic and the construction clearly satisfies the stabilizer outcome statistics for single shot experiments, as we effectively have reproduced the prepare-measure model given in the previous section.

Using the stabilizer measurement update rules described previously we can define the stochastic measurement update map to be $\Gamma_{k|M}$, letting $k = \gamma(M)$;

$$\Gamma_{k|M} : (P, \gamma) \mapsto (P', \gamma') \tag{5.11}$$

Such that;

Line 1: We set $P \mapsto P' = M$.

Note this update can be redefined to not " store" the previous measurement as follows: If $[P, M] \neq 0$ set $P' = M$, conversely if $[P, M] = 0$ set $P' = \{P, M, PM\}$ with equal probability. The update rule included in the paper was the simplest to state and therefore the one I chose to include. However the second update rule may be more in line with the quasi-probability simulation given in [69].

Line 2: If $[M, P] \neq 0$: $\left.\begin{array}{l} \gamma'(S) = \gamma(S) \\ \gamma'(SM) = \gamma(S) + k + \beta(S, M) \end{array}\right\} \forall S \in \widetilde{\mathcal{P}}_2, \, | \, [S, M] = [S, P] = 0.$

With knowledge of P we can only infer what the other possible stabilizer elements of the pre-measurement group could have been, namely any S such that $[S, P] = 0$. However with knowledge of M we can further reduce the set of stabilizer elements that could have been part of the post-measurement group, namely S such that $[S, M] = 0$. Therefore we update value of the phase function for any possible new element of the post-measurement stabilizer group SM, remembering that if S commutes with both P and M it's value should remain fixed through the measurement.

Line 3: if $[M, P] = 0$: $\gamma'(PM) = \gamma(P) + k + \beta(P, M)$.

If P and M commute then we can directly construct the post-measurement stabilizer group, given $P \neq M$ which implies a trivial update, which is $\mathcal{S}(\psi') = \{\mathbb{I}, (-1)^{\gamma(P)}P, (-1)^k M, (-1)^{\gamma(P)+k+\beta(P,M)}S_{P+M}\}$. Hence we need to update $\gamma'(PM)$ as given above.

Line 4: Set $\gamma'(S) = \{0, 1\}$ w. eq. prob $\forall S \in \widetilde{\mathcal{P}}_2, \, | \, [S, M] \neq 0, \, [S, P] = 0.$

This line randomizes all elements of the phase function that could have been part of the pre-measurement stabilizer group, but not the post-measurement stabilizer group. Ensuring no biasing of future measurement outcomes.

Line 5: Set $\gamma'(S) = \gamma(S)$ Otherwise.

This final line captures the fact that if none of the conditions above are met we must keep the value of the phase function static.

> **Lemma 5.5.1** *The ontological model presented above is a ψ-epistemic model of the 2-qubit stabilizer subtheory.*

Proof. The proof of correctness is subsumed by the general n-qubit proof, theorem 5.6.1. The direct proof of this lemma has been included in the appendix. We have included the proof as it gives an intuitive understanding of the structure of update rules in the model.

∎

We have also provided a mathematica code base[5] to simulate the 2-qubit model.

> **Note 5.2** *This code and an explanation of the functionality has been provided in appendix A.1.* ∎

5.5.1 Features of the 2-qubit Model

The Mermin-Peres Square

The Mermin-Peres (MP) square [55, 63, 38] is a well known proof of state-independent traditional contextuality. The proof only requires stabilizer measurement, therefore the n-qubit stabilizer formalism, for $n > 1$, is traditionally contextual. There are several variations of the MP square for 2-qubits. Here we will focus on demonstrating how the model reproduces the quantum statistics for the most common form of the square, figure 5.1.

---

[5]Accessible on github, date last accessed April 7th 2019.

$$
\begin{array}{ccccccc}
X\mathbb{I} & \text{---} & \mathbb{I}X & \text{---} & XX & \rightarrow & \mathbb{I} \\
| & & | & & | & & \\
\mathbb{I}Z & \text{---} & Z\mathbb{I} & \text{---} & ZZ & \rightarrow & \mathbb{I} \\
| & & | & & | & & \\
XZ & \text{---} & ZX & \text{---} & YY & \rightarrow & \mathbb{I} \\
\downarrow & & \downarrow & & \downarrow & & \\
\mathbb{I} & & \mathbb{I} & & -\mathbb{I} & &
\end{array}
$$

Figure 5.1: A Mermin-Peres square for 2-qubits, with tensor notation suppressed. Each row and column constitutes a set of commuting observables, who's product is given at the end of the arrows.

Under the assumptions of traditional non-contextuality there is no value assignment that satisfies the functional relationships between all commuting observables in the Mermin-Peres square. i.e. $\nexists v_\lambda \in \{+1, -1\}^{|\widetilde{\mathcal{P}}_2|}$ such that;

$$
\begin{aligned}
v_\lambda(X\mathbb{I})v_\lambda(\mathbb{I}X)v_\lambda(XX) &= +1, \\
v_\lambda(Z\mathbb{I})v_\lambda(\mathbb{I}Z)v_\lambda(ZZ) &= +1, \\
v_\lambda(X\mathbb{I})v_\lambda(\mathbb{I}Z)v_\lambda(XZ) &= +1, \\
v_\lambda(Z\mathbb{I})v_\lambda(\mathbb{I}X)v_\lambda(ZX) &= +1, \\
v_\lambda(XZ)v_\lambda(ZX)v_\lambda(YY) &= +1, \\
v_\lambda(XX)v_\lambda(ZZ)v_\lambda(YY) &= -1.
\end{aligned}
$$

The model here evades such a contradiction by implicitly demanding the value assignment is updated after any measurement. Further, we'd argue that a measurement update rule that updates value assignments is physically well justified: It is perfectly reasonable to demand that a measurement that extracts information from a system disturbs the state of the system, regardless of whether the measurement was performed with other commuting measurements. By inspecting the measurement update rules we can also see that we can only ever learn n-bits of information about the value assignment, due to the randomization step, satisfying the Holevo bound [37], similar to previously constructed contextual models of the Mermin-Peres square by Larsson and Kleinmann $et\ al.$ [47, 41].

To demonstrate how the model correctly reproduces the statistics of the Mermin-Peres square let us consider a test case of the $|00\rangle$ state, to be measured in one of the contexts. The model gives a deterministic outcome assignment to every Pauli observable, i.e. $v_\lambda(M) = (-1)^{\gamma_\lambda(M)}$. However, as we must update the ontic state after every measurement we can

113

only effectively learn the values assigned to one context. Suppose the ontic state of the system is $\lambda = (Z\mathbb{I}, \vec{0}) \in \operatorname{supp}(\mu_{00})$. We therefore have a value assignment to the observables in the Mermin-Peres square which can be expressed as;

$$
\begin{array}{ccc}
v_\lambda(M) = k & & v(M|\forall \lambda \in \operatorname{supp}(\mu_{00})) \\
\begin{array}{ccc} + & + & + \\ + & + & + \\ + & + & + \end{array}
&
\longleftrightarrow
&
\begin{array}{ccc} \{\pm\} & \{\pm\} & \{\pm\} \\ + & + & + \\ \{\pm\} & \{\pm\} & \{\pm\} \end{array}
\end{array}
$$

Where the second Mermin-Peres square has been used to demonstrate that cells not in the stabilizer group have uniformly random values. In the left-hand Mermin-Peres square the last column's constraint is not satisfied. So let us measure this problem context, starting with YY. To output an outcome of a measurement we simply read out the value of the phase function for that observable. So in this case $k_{YY} = +$. To perform the update rule we note that $[Z\mathbb{I}, YY] \neq 0$ so we update $Z\mathbb{I} \to YY$ and follow the non-commuting cases in the update rules. The post-measurement update to the phase function is given by line 2 of equation 5.11, so $\gamma'(S) = \gamma(S)$ and $\gamma'(SM) = \gamma(S) + k + \beta(S, M)$, $\forall S \in \widetilde{\mathcal{P}}_2$ such that $[S, YY] = [S, Z\mathbb{I}] = 0$. So we need to find the set of 2 qubit Pauli operators that commute with both $Z\mathbb{I}$ and YY. Namely $\{\mathbb{I}, ZZ, \mathbb{I}Y, ZX\}$ which forms a non-abelian subgroup of the Pauli group[6]. Following the phase function update rules we need to update phase function to:

$$
\begin{aligned}
\gamma'(YY) &= \gamma(YY) = k = 0, \\
\gamma'(ZZ) &= \gamma(ZZ) = 0, \\
\gamma'(XX) &= \gamma(ZZ) + k + \beta(ZZ, YY) = 1, \\
\gamma'(ZX) &= \gamma(ZX) = 0, \\
\gamma'(XZ) &= \gamma(ZX) + k + \beta(ZX, YY) = 0.
\end{aligned}
$$

Finally we randomize any elements that commute with $Z\mathbb{I}$ and anti-commute with YY, i.e. in this Mermin-Peres sqaure $\mathbb{Z}\mathbb{I}$, $\mathbb{I}Z$, and $\mathbb{I}X$. This means, after converting the phase function, the value assignment for the Mermin-Peres square after a YY measurement is now, chosing the randomization to give $+$;

$$
\begin{array}{ccc}
v(M|\lambda') = k & & v(M|\forall \lambda' \in \operatorname{supp}(\mu_{B_{11}})) \\
\begin{array}{ccc} + & + & - \\ + & + & + \\ + & + & + \end{array}
&
\longleftrightarrow
&
\begin{array}{ccc} \{\pm\} & \{\pm\} & - \\ \{\pm\} & \{\pm\} & + \\ \{\pm\} & \{\pm\} & + \end{array}
\end{array}
$$

---

[6]Note that the Mermin-Peres square gives graphical way to find such operators.

Where we have colored the deterministically updated cells blue and the randomized cells green. Note the final row's value assignments are now correlated but random, this makes sense when we consider that given $P = Z\mathbb{I}$ the pre-measurement stabilizer group could have been $\{\mathbb{I}, Z\mathbb{I}, \mathbb{I}Z, ZZ\}$ or $\{\mathbb{I}, Z\mathbb{I}, \mathbb{I}X, ZX\}$, so the update rule must account for both possibilities.

The update rule has enforced that the contexts that YY is an element of satisfy the functional relationships, as the new ontic state is in the support of any state stabilized by YY. Further, as expected, the value assignments are consistent with the new stabilizer group $S(B_{11}) = \{\mathbb{I}, -XX, ZZ, YY\}$, where $|B_{11}\rangle = \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle)$. If we were to go through the update rules for every other measurement in the context (the third column) we would see that the value assignment for the square only performs the randomization step.

The PBR theorem

One obstacle to constructing any ontological model of the stabilizer formalism is the Pusey-Barrett-Rudolph no-go theorem [66]. This no-go result states that any ontological model that treats preparations on subsystems independently of preparations on other subsystems, referred to as *preparation independence*, must be ψ-ontic.

As the model presented here is ψ-epistemic the PBR theorem implies the model must represent preparations non-locally. This indeed is the case. For example consider the $|00\rangle$ preparation, the stabilizer group of $|00\rangle$ is given by $\mathcal{S}(|00\rangle) = \{\mathbb{I}, Z\mathbb{I}, \mathbb{I}Z, ZZ\}$ and therefore the support of $\mu_{|00\rangle}$ only contains ontic states that satisfy $\gamma(Z\mathbb{I}) + \gamma(\mathbb{I}Z) = \gamma(ZZ)$, however the same cannot be said for other local variables such as $\gamma(XI) + \gamma(IZ) \neq \gamma(XZ)$. Hence the model evades the PBR theorem by being explicitly non-local, even for product preparations.

The anti-distinguishing measurements used in the PBR theorem also place additional constraints on the structure of possible ψ-epistemic models, as demonstrated by Karanjai *et al.*. Namely they imply that any set of states that can be anti-distinguished cannot have a global intersection[7] [40]. A quick sketch of the proof is as follows [50]; Suppose we have a set of states $S = \{\rho_i\}_i$ that can be antidistinguished by some POVM, with elements E_j.

---

[7]Karanjai *et al.* also proved strict lower bound on the size of the ontic space of any ontological model of the stabilizer formalism, which the model presented here satisfies. This is covered in more detail in the appendix.
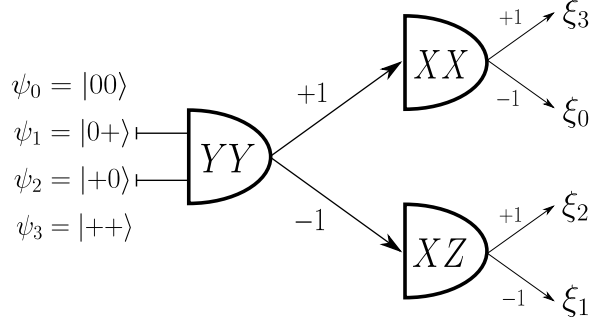
Figure 5.2: The Pusey-Barrett-Rudolph anti-distinguishing adaptive measurement circuit. The input to the circuit is a ψ_i, which all give a random outcome for the measurement of YY. The second measurement is then conditioned on the outcome of YY. The possible sets of outcomes then anti-distinguishes the preparation. For example $\mathcal{O}(YY) = +1$ and $\mathcal{O}(XX) = -1$ implies ψ_0 wasn't prepared, hence this path is labelled ξ_0.

So;

$$\mathrm{Tr}(\rho_i E_j) \begin{cases} = 0, & \text{if } i = j, \\ \geq 0, & \text{otherwise.} \end{cases}$$

I.e. if we measured outcome j we know with certainty that ρ_j was not prepared. This implies that in any ontological model that reproduces the statistics of an anti-distinguishing measurement $\xi_j(\lambda) = 0$, $\forall \lambda \in \mathrm{supp}(\mu_{\rho_j}) = \Delta_{\rho_j}$, as $\mathrm{Tr}(\rho_j E_j) = \int_\Lambda d\lambda \, \xi_j(\lambda) \mu_{\rho_j}(\lambda) = \int_{\Delta_{\rho_j}} d\lambda \, \xi_j(\lambda) \mu_{\rho_j}(\lambda) = 0$. If we assume that μ_{ρ_i} share a joint support, $\exists \Delta \subset \Delta_{\rho_i}$, $\forall i$, then the same argument holds for all outcomes, i.e. $\xi_i(\lambda) = 0$, $\forall \lambda \in \Delta$, $\forall i$. However by normalization we require $\sum_i \xi_i(\lambda) = 1$, $\forall \lambda$, therefore we have a contradiction.

While these anti-distinguishing POVMs cannot be measured by a single measurement in the stabilizer formalism, we can measure them via a sequence of adaptive measurements. To demonstrate how the model accounts for these sequences of adaptive measurements we use the anti-distinguishable set of states $\{|00\rangle, |0+\rangle, |+0\rangle, |++\rangle\}$ from the PBR paper [66]. The adaptive sequence of stabilizer measurements that anti-distinguishes this set of states is given in figure 5.2.

The measurement of the first observable, with outcome k, updates the stabilizer groups

of the input states as follows;

$$\begin{array}{ll}
|00\rangle & \equiv \{\mathbb{I}, \mathbb{I}Z, Z\mathbb{I}, ZZ\} \\
|0+\rangle & \equiv \{\mathbb{I}, \mathbb{I}X, Z\mathbb{I}, ZX\} \\
|+0\rangle & \equiv \{\mathbb{I}, \mathbb{I}Z, X\mathbb{I}, XZ\} \\
|++\rangle & \equiv \{\mathbb{I}, \mathbb{I}X, X\mathbb{I}, XX\}
\end{array} \xrightarrow{\mathcal{O}(YY)=k}
\begin{array}{l}
\left\{\mathbb{I}, ZZ, (-1)^k YY, (-1)^{k+1} XX\right\}, \\
\left\{\mathbb{I}, ZX, (-1)^k YY, (-1)^k XZ\right\}, \\
\left\{\mathbb{I}, XZ, (-1)^k YY, (-1)^k ZX\right\}, \\
\left\{\mathbb{I}, XX, (-1)^k YY, (-1)^{k+1} ZZ\right\}.
\end{array} \qquad (5.12)$$

By noting that two stabilizer states, ρ_1 and ρ_2, are orthogonal if and only if $\exists P \in \mathcal{S}(\rho_1)$ and $-P \in \mathcal{S}(\rho_2)$. We can see that the first and last post-measurement stabilizer states are orthogonal if $k = 0$ and equivalent if $k = 1$. Similarly, the second and third post-measurement groups are orthogonal if $k = 1$ and equivalent otherwise. Therefore by adaptively measuring XX or XZ, equally ZZ or ZX, after the first measurement we can rule out a single preparation, anti-distinguishing the set of input states.

The first thing to note about how the model reproduces the PBR statistics is that it satisfies the no joint overlap requirement. By investigating the stabilizer groups of the pre-measurement states above it is clear that the 4 states do not share a mutual generator and therefore have no joint overlap. Mathematically, $\nexists P$ such that $(P, \cdot) \in \text{supp}(\mu_{\psi_i})$, $\forall i$ therefore $\cap_i \text{supp}(\mu_{\psi_i}) = \emptyset$.

To demonstrate how the measurement update rule gives the correct anti-distinguishing measurement statistics we investigate the evolution of the representations of $|00\rangle$ and $|++\rangle$, a similar analysis can be given for $|0+\rangle$ and $|+0\rangle$. Firstly we let $\gamma(YY) = k$ for both states. Considering that the $\lambda = (P, \cdot)$ in the support of $\mu_{|00\rangle}$ and $\mu_{|++\rangle}$ are $P \in \{\mathbb{I}Z, Z\mathbb{I}, ZZ\}$ and $P \in \{\mathbb{I}X, X\mathbb{I}, XX\}$, respectively, therefore $\mu_{|00\rangle} \cap \mu_{|++\rangle} = \emptyset$. For completeness, we examine both $[P, YY] = 0$ and $[P, YY] \neq 0$.

Taking our input state to be $\psi = |00\rangle$ and $(P, \cdot) = (\mathbb{I}Z, \cdot)$ the update rules determine that we must update all $S \in \widetilde{\mathcal{P}}_2 \backslash \mathbb{I}$ such that $[S, YY] = [S, \mathbb{I}Z] = 0$, as $[\mathbb{I}Z, YY] \neq 0$. This set is $\{\mathbb{I}, ZZ, XZ, Y\mathbb{I}\}$, therefore we perform the update;

$$\begin{aligned}
\gamma'(ZZ) &= \gamma(ZZ) = 0, \\
\gamma'(XX) &= \gamma(ZZ) + k + \beta(ZZ, YY) = k + 1,
\end{aligned}$$

as required by equation 5.12. Note we also update $\gamma'(XZ) = \gamma(XZ)$ and $\gamma'(ZX) = \gamma(XZ) + k + \beta(XZ, YY) = \gamma(XZ) + k$, which as $\gamma(XZ)$ is random implies $\gamma'(ZX)$ is random. This update is important because the ontic state we have considered is in the support of $|+0\rangle$, therefore if $\gamma(XZ)$ is fixed by consistency it gives the correct update for that state. Applying the above to $(\mathbb{I}X, \cdot) \in \text{supp}(\mu_{|++\rangle})$, where now the set of commuting

operators is $\{\mathbb{I}, XX, ZX, Y\mathbb{I}\}$, we have;

$$\gamma'(XX) = \gamma(XX) = 0,$$
$$\gamma'(ZZ) = \gamma(XX) + k + \beta(XX, YY) = k + 1,$$

as required.

By symmetry we can infer the update rules for $\lambda_1 = (Z\mathbb{I}, \cdot)$ and $\lambda_2 = (X\mathbb{I}, \cdot)$ will follow a similar structure. So finally we must check the commuting cases $[ZZ, YY] = [XX, YY] = 0$. By the update rules we know in this case we leave $\gamma(P)$ static and update $\gamma'(PM)$, therefore given $\lambda_1 = (ZZ, \cdot)$ we perform the updates $\gamma'(ZZ) = \gamma(ZZ)$ and $\gamma'(XX) = \gamma(ZZ) + k + \beta(ZZ, YY) = k + 1$ and for similarly for $\lambda_2 = (XX, \cdot)$ we have $\gamma'(XX) = \gamma(XX)$ and $\gamma'(ZZ) = \gamma(XX) + k + \beta(XX, YY) = k + 1$, as required.

Traditional Contextuality

In the ontological model presented in this paper the phase function clearly defines a value assignment to all Pauli observables. However in comparison to the Mermin style proofs of contextuality [55, 63, 56] we do not assume that the value assignments of all commuting observables can be simultaneously extracted. Further any measurement procedure necessarily disturbs the value assignments of commuting observables. Therefore as expected the model is Kochen-Specker contextual.

However we argue that the contextuality present in the stabilizer formalism is not a surprising form of contextuality. Taking the Mermin-Peres square as our example, there is no reason to believe that the outcome of YY should be the same independent of the context it is measured in. Nor should its outcome be the same independent of where in the sequence of measurements it is performed. The only requirement imposed by the stabilizer formalism is that we satisfy the algebraic constraints, given by the stabilizer group operation, on observables in a stabilizer group.

To illustrate this let us consider four measurement sequences $S_1 = XX \to ZZ \to YY$, $S_2 = XZ \to ZX \to YY$, $S_1' = YY \to ZZ \to XX$, and $S_2' = YY \to ZX \to XZ$. The first thing to note is that the first two measurements in each sequence define a rank-1 projective measurement, which fixes the outcome of the final measurement. For the first sequence we have the constraint $S_1 : \mathcal{O}(XX)\mathcal{O}(ZZ) \oplus 1 = \mathcal{O}(YY)$, similarly $S_2 : \mathcal{O}(XZ)\mathcal{O}(ZX) = \mathcal{O}(YY)$. The usual contradiction is derived by adding in contexts that jointly constrain the outcomes of XX, ZZ, XZ, and ZX. However this ignores what the algebraic constraints are imposing, i.e. the outcome of a YY measurement is constrained by the outcomes of

118

the previously measured observables. Not on the possible, counter-factual, measurement sequences.

The final two sequences, S_1' and S_2', illustrate that such an update to the value assignment is physically justified. Consider measuring a context in the Mermin-Peres square with the maximally mixed state as the input. In this case the outcome of a YY measurement should be random, by lemma 5.3.1. If we measure sequence S_1' and S_2' the outcome $\mathcal{O}(YY)$ is clearly random, by definition of the phase function. Similarly if we measure sequences S_1 or S_2 the outcome $\mathcal{O}(YY)$ is still random, as the outcomes of the previous measurements in the sequence are random, but not necessarily equal to the outcome if YY was measured first. Therefore there is no reason to expect the outcome of YY to be the same regardless of whether it was measured first or last, as the only constraint we should obey is that it's outcome is random. Hence, the only case where we should expect a fixed outcome for YY is if the input state was an eigenstate of YY, in which case the algebraic constraints limit our value assignments to other commuting observables, but crucially no others.

5.6 Generalizing the model to n-qubits

To begin we first establish some notation, denote a proper abelian subgroup of the projective Pauli group with bold font, i.e. $\boldsymbol{G} \subset \widetilde{\mathcal{P}}_n$. We define the set of all Abelian subgroups with k generators as $\boldsymbol{\mathcal{P}}_k := \left\{ \boldsymbol{G} \mid \boldsymbol{G} \subset \widetilde{\mathcal{P}}_n, \, |\boldsymbol{G}| = 2^k \right\}$. These subgroups will replace the role of $\widetilde{\mathcal{P}}_2 \backslash \mathbb{I}$ in the two qubit model, and technically for the two qubit model the stabilizer ontology is $\boldsymbol{\mathcal{P}}_1$.

The ontology of the n-qubit model is defined to be $\Lambda = \boldsymbol{\mathcal{P}}_{n-1} \times \gamma_n$. The uniform distributions representing n-qubit pure stabilizer states $\psi \in \mathcal{S}(\mathcal{H}_{2^n})$ have support;

$$\mathrm{supp}(\mu_\psi) = \left\{ \lambda = (\boldsymbol{G}, \gamma) \mid \boldsymbol{G} \subset \widetilde{\mathcal{S}}(\psi), \, \boldsymbol{G} \in \boldsymbol{\mathcal{P}}_{n-1}, \, \gamma \cong \psi \right\}. \tag{5.13}$$

These states, while being defined on an exponentially growing ontology, are ψ-epistemic. For example the ontic state $\lambda = (\langle Z_i \rangle_{i=1,..,n-1}, \gamma = \vec{0})$ is in the support of $|0\rangle^{\otimes n}$ and $|0\rangle^{\otimes n-1} |+\rangle$. To find which pure stabilizer states have support on a given ontic state note that given a $\boldsymbol{G} \in \boldsymbol{\mathcal{P}}_{n-1}$ there is a finite set of non-commuting Pauli operators that commute with $\boldsymbol{G}$ that can all be added to the group to construct valid stabilizer groups.

As before the Clifford operations in the model can be inferred from their action of the projective Pauli group. So a Clifford operation's, C, ontological representation, Γ_C, is

given by;

$$\Gamma_C : (\boldsymbol{G}, \gamma) \mapsto (C\boldsymbol{G}C^\dagger, \gamma^R + \gamma_c), \tag{5.14}$$

Where again we re-order γ to match the Clifford permutation and the conjugation is take to be on every element of $\boldsymbol{G}$.

As in the two qubit model the response functions read-out the value of the phase function;

$$\xi_{k,M}(\lambda) = \begin{cases} 1 & \text{if } \gamma(M) = k, \\ 0 & \text{otherwise.} \end{cases} \tag{5.15}$$

> **Note 5.3** *As with the two-qubit update rules I have presented a more in-depth analysis of the n-qubit update rules compared to the published version of this paper.* ∎

We define the measurement update rules for a measurement of M with outcome k to be;

$$\Gamma_{k|M} : (\boldsymbol{G}, \gamma) \mapsto (\boldsymbol{G}', \gamma') \tag{5.16}$$

Such that;

Line 1: We set;

$$\boldsymbol{G} \mapsto \boldsymbol{G}' = \langle \boldsymbol{G}_M, M \rangle \text{ w. eq. prob } \boldsymbol{G}_M \in \boldsymbol{\mathcal{P}}_{n-2} \,|\, [M, G_i] = 0, \, \boldsymbol{G}_M \subset \boldsymbol{G}, \, \forall G_i \in \boldsymbol{G}_M.$$

Note this update can be redefined to not " store" the previous measurement as follows: If $[\boldsymbol{G}, M] \neq 0$ set $\boldsymbol{G}' = \langle \boldsymbol{G}_M, M \rangle$. Conversely if $[\boldsymbol{G}, M] \neq 0$ uniformly sample a subgroup $\boldsymbol{H} \in \boldsymbol{\mathcal{P}}_{n-1}$ from $\langle \boldsymbol{G}, M \rangle$ and set $\boldsymbol{G}' = \boldsymbol{H}$. Again update rule included in the paper was the simplest to state and therefore the one I chose to include. However again the second update rule may be more in line with the quasi-probability simulation given in [69].

Line 2: If $[M, \boldsymbol{G}] \neq 0$:

$$\left.\begin{array}{l} \gamma'(S) = \gamma(S) \\ \gamma'(SM) = \gamma(S) + k + \beta(S, M) \end{array}\right\} \,\forall S \in \widetilde{\mathcal{P}}_n, \,|\, [S, M] = [S, \boldsymbol{G}] = 0.$$

120

If M does not commute with $\boldsymbol{G}$ then we must perform an update similar to the 2-qubit case. We find all S that could have been an element of the pre and post measurement groups and update the phase function for the possible new elements of the group SM. See the proof of theorem 5.6.1 for a full explanation of how this correctly updates the phase function.

Line 3: If $[M, \boldsymbol{G}] = 0$:

$$\gamma'(G_i M) = \gamma(G_i) + k + \beta(G_i, M) \ \forall G_i \in \boldsymbol{G}.$$

If M commutes with all of $\boldsymbol{G}$ then we update the phase function for all new elements the post-measurement stabilizer group $\{G_i M | \forall G_i \in \boldsymbol{G}\}$.

Line 4: Set:

$$\gamma'(S) = \{0, 1\} \text{ w. eq. prob} \, \forall S \in \widetilde{\mathcal{P}}_n, \, | \, [S, M] \neq 0, \, [S, G_i] = 0, \, \forall G_i \in \boldsymbol{G}.$$

This line randomizes all elements of the phase function that could have been part of the pre-measurement stabilizer group, but not the post-measurement stabilizer group. Ensuring no biasing of future measurement outcomes.

Line 5: Set $\gamma'(S) = \gamma(S)$ Otherwise.

This final line captures the fact that if none of the conditions above are met we must keep the value of the phase function static.

> **Theorem 5.6.1** *The ontological model presented above is a ψ-epistemic model of the n-qubit stabilizer subtheory.*

Proof. The model is ψ-epistemic by the example given above. It also reproduces the single shot statistics, by the argument given for the prepare-measure-discard model. Therefore we just to check that the measurement update rule maps the phase function to a new phase function that is consistent with the post-measurement state.

The first line corresponds to updating $\boldsymbol{G}$ to a new subgroup that is an proper subgroup of the post-measurement state. As there are many possible proper subgroups of $\boldsymbol{G}$ that commute with M if $[M, G_i] = 0, \, \forall G_i \in \boldsymbol{G}$ we sample one of them uniformly. If M does not commute with all elements of $\boldsymbol{G}$ there is only one such $\boldsymbol{G}_M$. $\langle \boldsymbol{G}_M, M \rangle$ is an proper subgroup of the post-measurement group and therefore $(\boldsymbol{G}', \cdot) \in \text{supp}\mu_{\rho'}$

We will prove correctness by running through each of the possible measurement updates that need to be applied, which are categorized by the commutation relation between $\boldsymbol{G}$ and M.

The first case we consider is $[M, \boldsymbol{G}] \neq 0$, i.e. M does not commute with 2^{n-2} elements of $\boldsymbol{G}$. We can write the post-measurement group as $\langle \boldsymbol{G}_M, M, S \rangle$, where S is the unknown stabilizer element, such that $[S, M] = [S, \boldsymbol{G}] = 0$, $S \notin \boldsymbol{G}$ and $\boldsymbol{G}_M$ is the subgroup of $\boldsymbol{G}$ containing all elements that commute with M. There are 3 non-trivial equivalence classes of S, which mutually anti-commute. Any $S \in \widetilde{\mathcal{P}}_n$ that commutes with both M and $\boldsymbol{G}$ could have been elements of the pre-measurement and post-measurement group, implying their phase function should remain fixed through the measurement. We therefore update all possible new elements of the post-measurement group, which are $SM \notin \mathcal{S}(\rho)$ as $\exists G \in \boldsymbol{G}$ such that $[SM, G] \neq 0$. Given two possible extensions generated by S_1 and S_2 such that $[S_1, S_2] \neq 0$ we have $[S_1 Mg, S_2 Mg'] \neq 0, \forall g, g' \in \boldsymbol{G}_M$. This implies the post measurement equivalence classes do not intersect, i.e. $S_1 Mg \notin \langle \boldsymbol{G}_M, M, S_2 \rangle, \forall g \in G_M$. Therefore SM is only an element of one possible group extension implying that $\gamma'(SM)$ has the correct phase if $SM \in \mathcal{S}(\rho')$. Note that any $G \in \boldsymbol{G}$ that commutes with M satisfies the conditional of lines 2 and 3. Finally any element that could have been part of the pre-measurement group, but not the post-measurement group is randomized, line 4.

The second case we need to consider is when M commutes with all elements of $\boldsymbol{G}$. Again this can be broken down to two cases $M \notin \boldsymbol{G}$ and $M \in \boldsymbol{G}$. If M commutes with all of $\boldsymbol{G}$ but is not an element of the group, then we know the post-measurement stabilizer group can be expressed as $\langle \boldsymbol{G}, (-1)^k M \rangle$. Therefore the new (or potentially old) elements of the group are $G_i M$ with phase $\gamma(G_i M) = \gamma(G_i) + k + \beta(G_i, M) = \gamma'(G_i M)$. Note if these Pauli operators were elements of the stabilizer group prior to measurement their phase remains unchanged. If $M \in \boldsymbol{G}$ then the phase function should remain unchanged through the measurement, up to randomization on stabilizer measurements not in the stabilizer group. This is expressed in lines 3 through 5 of the update rule. Note line 5 will be satisfied by operators that commute with both M and $\boldsymbol{G}$, but are not in $\boldsymbol{G}$, these are possible elements of the stabilizer group and therefore their phase function should remain fixed.

Therefore the update rule given above maps an ontic state $\lambda = (\boldsymbol{G}, \gamma) \in \operatorname{supp}(\mu_\psi)$ to an ontic state $\lambda' = (\boldsymbol{G}', \gamma') \in \operatorname{supp}(\mu_{\psi'})$, where $\rho_{\psi'} \propto \frac{\mathbb{I}+(-1)^k M}{2} \rho_\psi \frac{\mathbb{I}+(-1)^k M}{2}$. Further the randomization procedure ensures a random value of the phase function for observables not in the post-measurement stabilizer group. Therefore the model correctly reproduces the stabilizer statistics.

∎

It is interesting to note that the additional stabilizer ontology we introduced has not been required prior to defining the update rules. However, we can actual use this ontology to streamline the model to define an always-ψ-epistemic model. We can achieve this by noticing that when $[\boldsymbol{G}, M] \neq 0$ the outcome of the measurement should be random, and therefore we can reconstruct the response functions accordingly. An overview of this procedure is covered in the appendix.

The generalization of the model to n-qubits has required us to extend the stabilizer ontology to proper subgroups of maximal abelian subgroups of the Stabilizer operators. It is not clear whether this additional ontology is required or can be made more compact. However by reducing the size of the groups in the ontology leads to ambiguities in how the phase function should be updated during measurement. It is therefore a highly interesting open problem whether $n-1$ generators of a stabilizer state are truly required to correctly update an epistemic state in an arbitrary ψ-epistemic model of the n-qubit stabilizer formalism. If this is indeed the case, it would suggest *almost* full knowledge of a stabilizer state[8] is required to be encoded in the ontology to correctly reproduce the stabilizer subtheory's statistics.

5.6.1   The n-qubit model's relation to the single-qubit 8-state model

Corollary 5.6.2 *The $n = 1$ case of the model is equivalent to the 8-state model.*

Proof. To see this let us consider the ontology for a single qubit in the model is $\Lambda = \boldsymbol{P}_0 \times \gamma_{n=1}$. As $\boldsymbol{P}_0 = \{\mathbb{I}\}$ and all stabilizer groups contain $\mathbb{I}$, the stabilizer group part of the ontology is trivial and we can remove it. The non-trivial ontology is therefore $\Lambda = \gamma_1$, which is identical to the 8 possible value assignments over X, Y, and Z in the 8-state model.

States are defined as having uniform support over all consistent phase functions. Writing the phase function as $\gamma = (0, x, y, z)$ we see that a stabilizer state $\rho = \frac{\mathbb{I}+(-1)^{s_P}P}{2}$ has support on all phase functions such that $\gamma(P) = s_P$. I.e. a stabilizer state has support on 4 ontic states as in the 8-state model, and the supports in coincide. Therefore stabilizer states in the model presented here and the 8-state model have the same representation.

Clifford transformations are represented as permutations of the ontic space in both our model and the 8-state model. In the 8-state model the permutation maps representing a

---

[8]As n generators uniquely specify the state.

Clifford operations are derived from their stabilizer relations, i.e. $\Gamma_{8-\text{state}}(H) : (x, y, z) \mapsto (z, y \oplus 1, x)$. This also holds in the model presented here, i.e. $\Gamma_{\text{model}}(H) : (x, y, z) \mapsto (z, y, x) \oplus (0, 1, 0)$, where we have decomposed the map into γ^R and γ_c for clarity.

The measurement response functions are defined identically, with the response function essentially reading out the value of the phase function/value assignment. All that is left is to demonstrate is that the update rules are the same. In the single-qubit stabilizer formalism all measurements are rank-1 projectors. Therefore the measurement update map in the 8-state model can be simply stated as repreparing the eigenstate corresponding the eigenvalue measured. For example an X measurement with a $+1$ outcome the measurement update rule prepares the distribution $\mu_{|+\rangle}$.

Returning to the $n = 1$ case of the updates rules for the model, equation 5.16. As all single qubit measurements commute with $\mathcal{P}_0$ ontology we only consider the commuting case of the update rules. In this case we have $\gamma'(G_i M) = \gamma(G_i) + k + \beta(G_i, M)$, $\forall G_i \in \mathbf{G}$, which for a single qubit reduces to setting $\gamma'(M) = k$. Finally we randomize all phase function elements that don't commute with M, which in the case of a single qubit Pauli observable is all other single qubit Pauli observables. This update effectively reprepares the eigenstate of the measured observable with eigenvalue k. Therefore the update rules are the same. This implies the 8-state model and the $n = 1$ case of the ontological model of the qubit stabilizer formalism presented here are the equivalent. ∎

5.7 Discussion

In this paper we have constructed a contextual ψ-epistemic model of the n-qubit stabilizer formalism. The model is constructed by considering that given a stabilizer state it is possible to reproduce the one-shot statistics of stabilizer measurements by uniform sampling over value assignments, which we term the phase function, that are consistent with a stabilizer state. Measurement update rules then ensure that this value assignment is mapped to a value assignment consistent with the post-measurement state. Therefore reproducing the statistics of the n-qubit stabilizer formalism.

The core component of the model's presentation is the value assignment. We have constructed the model this way to investigate the structure of contextuality in the n-qubit stabilizer formalism. However, to correctly update the value assignment through a measurement we are need to encode $n - 1$ generators of the stabilizer state in the ontology. Whether this additional ontology is required or can be reduced is an interesting open question.

By moving away from an outcome deterministic model it is possible to use the presented model to construct an *always-ψ-epistemic* model, covered in the appendix. Always-ψ-epistemic models have a strong resemblance to probabilistic classical models as a quantum state represents a true lack of knowledge of the physical state of a system. Therefore they present an interesting possibility of a categorizing the transition from classical to quantum theory. Further they share many properties with weak-simulations and could providing an intriguing formalism for investigate the resources required to simulate quantum systems, universal or otherwise.

5.8   Proof of correctness of the 2-qubit ontological model

Here we give the proof of lemma 5.5.1;

Proof. From the argument given for the prepare-measure-discard model we know that we will correctly reproduce the stabilizer outcome statistics if the phase function satisfies $\gamma(M) = p_b$, $\forall b \in \mathcal{B}(\mathcal{S}(\rho))$ and is random otherwise. The distributions representing preparations clearly satisfy this, therefore we just need to show the update rule maps γ to a γ' such that $\gamma'(M) = p_{b'}$, $\forall b' \in \mathcal{B}(\mathcal{S}(\rho'))$ and random otherwise. Additionally by the stabilizer update rule we know that $(-1)^k M \in \mathcal{S}(\rho')$ so the update on the stabilizer operator portion of the ontology is in the support of the post-measurement state.

To prove the phase function update map correct updates to a phase function consistent with the post-measurement state we will exhaustively prove each possibility, with each case defined via the commutation relation between M and P.

For the first case, we let $[P, M] \neq 0$, i.e. M was certainly not an element of the pre-measurement stabilizer group. We can always write this group as;

$$\left\{ \mathbb{I}, (-1)^{\gamma(P)} P, (-1)^{\gamma(S)} S, (-1)^{\gamma(S)+\gamma(P)} PS \right\},$$

where $[M, PS] = [M, P] \neq 0$ and $[S, M] = 0$. Therefore the post measurement group is given by;

$$\left\{ \mathbb{I}, (-1)^{\gamma(S)} S, (-1)^k M, (-1)^{\gamma(S)+k+\beta(S,M)} P_{S+M} \right\}.$$

125

I.e. we need to keep $\gamma'(S) = \gamma(S)$ static through the update and update the phase function on P_{S+M} to $\gamma'(SM) = \gamma(S) + k + \beta(S, M)$. This is captured by lines 2 and 3 of the update rule. However in the update rule we search over all such S, we do this because given a P there are 3 possible non-trivial stabilizer groups P is an element of, modulo phases. I.e. there are 3 non-commuting Pauli operators, $\{S_1, S_2, S_3\}$ we could combine with P to construct all possible maximal stabilizer groups containing P. These groups only share the elements $\mathbb{I}$ and P, this can easily be verified by checking PS_i's commutation relations with $S_{j \neq i}$ and $PS_{j \neq i}$, therefore by updating all possible S we cover all possible pre-measurement stabilizer groups. Further by the consistency of the phase function with the pre-measurement state we know if $S \in \mathcal{S}(\rho)$ then $\gamma(S) = p_S$, and if not $\gamma(S)$ is a random value. Therefore the phase function remains random on elements not in the group. Finally for any S such that $[P, S] = 0$ and $[M, S] \neq 0$ we randomize $\gamma(S)$ as S could have been an element of the pre-measurement group, but is definitely not an element after. Therefore γ' is consistent with all possible post-measurement stabilizer groups.

The second case is that $[M, P] = 0$ and $M \neq P$. Regardless of what the pre-measurement group was we can actually directly construct the post-measurement group;

$$\left\{ \mathbb{I}, (-1)^{\gamma(P)} P, (-1)^k M, (-1)^{\gamma(P)+k+\beta(P,M)} P_{P+M} \right\}.$$

Therefore we only need to update $\gamma'(PM) = \gamma(P) + k + \beta(P, M)$ and randomize all phase function elements that could have been elements of the pre-measurement group, but cannot be part of the post-measurement group, lines 4 and 5 of the update rule. Finally we have the case that $P = M$. This case is trivial and no update rule should be applied as pre-and-post measurement stabilizer groups are the same. However this case is rolled into the previous case in the update rule above, this can be seen by noting $\gamma'(PM) = \gamma(MM) = \gamma(\mathbb{I}) = 0$ and the randomizing step does not effect any potential stabilizer elements.

Therefore the update rule $\Gamma_{k|M}$ maps a phase function γ to a set of phase functions $\{\gamma'_i\}_i$ such that $\gamma'_i(P_{b'}) = p_{b'}$, $\forall i$, $\forall b \in \mathcal{B}(\mathcal{S}(\rho'))$ and $\gamma'_i(M)$ is uniformly random over i if $\pm P_{b'} \notin \mathcal{S}(\rho')$. And therefore reproduces the stabilizer statistics for sequential measurements.

∎

5.9 Streamlining the model

In the main body of this paper we have presented the ontological model of the stabilizer formalism in such a way that the phase function encodes the stabilizer statistics. This

has been done for ease of presentation. However, we can reduce the size of the model by utilizing the stabilizer part of the ontology.

Recall the ontology of the model is defined to be $\Lambda_n = \boldsymbol{\mathcal{P}}_{n-1} \times \gamma_n$. This definition of the ontology of the model grows super-exponentially;

$$|\Lambda_n| = |\boldsymbol{\mathcal{P}}_{n-1}| \times |\gamma_n| = \left[2^{(n-1)(n-2)/2} \prod_{k=0}^{n-2} (4^n - 1)\right] \times 2^{4^n},$$

where we have used a similar argument as [2] for the first term. Clearly 2^{4^n} dominates for all but the smallest n. Therefore the ontology grows super-exponentially with number of qubits, making it a poor candidate for a physically motivated ontological model.

The super-exponential scaling derives from us allowing all possible value assignments the model. However if consider the stabilizer statistics, lemma 5.3.1, we see that $\exists S \in \mathcal{S}(\rho)$ such that $[S, M] \neq 0$ then the outcome of a measurement of M should be random.

Therefore if we have an ontic state $\lambda = (\boldsymbol{G}, \gamma)$ then we know the outcome of any measurement M will be random if $[\boldsymbol{G}, M] \neq 0$. Further this randomness can be encoded in the response functions, rather than the phase function;

$$\xi_{k,M}(\lambda) = \begin{cases} 1 & \text{if } \gamma(M) = k, \ [\boldsymbol{G}, M] = 0 \\ \frac{1}{2} & \text{if } [\boldsymbol{G}, M] \neq 0 \\ 0 & \text{otherwise.} \end{cases}$$

Therefore we do not need to store the value of the phase function if a Pauli operator does not commute with $\boldsymbol{G}$.

Further as we know the phase function must be consistent with a stabilizer state we can infer that we only need to store the phase on a set of generators of $\boldsymbol{G}$, due to the group structure of the stabilizer groups. I.e. it must be the case that $\gamma(g_1 + g_2) = \gamma(g_1) + \gamma(g_2) + \beta(g_1, g_2), \forall g_1, g_2 \in \boldsymbol{G}$.

Finally as $\boldsymbol{G} \in \boldsymbol{\mathcal{P}}_{n-1}$ there are only three possible equivalence classes of generators we can add to $\boldsymbol{G}$ to construct a pure stabilizer state's stabilizer group. Denote these possible extensions as $\mathcal{S}(\psi_1) = \langle \boldsymbol{G}, S_1 \rangle$, $\mathcal{S}(\psi_2) = \langle \boldsymbol{G}, S_2 \rangle$, and $\mathcal{S}(\psi_3) = \langle \boldsymbol{G}, S_3 \rangle$, where $\mathcal{S}(\psi_1) \neq \mathcal{S}(\psi_2) \neq \mathcal{S}(\psi_3)$. As the phase function must be consistent with one of these three group extensions, we only need to store the phase of the relevant generator, $\gamma(S_i)$. However if we wish not to be able to infer the quantum state from the ontic state we should treat each extension equivalently. So we demand the phase function also satisfies

$\gamma(S_i+g) = \gamma(S_i)+\gamma(g)+\beta(S_i, g), \forall S_i$, which is satisfiable via the non-intersecting property of $\{\mathcal{S}(\psi_i)\backslash\boldsymbol{G}\}_i$.

Therefore can construct an *symmetric-always-ψ-epistemic* model by encoding the ontic state in a tableau similar to the Gottesman-Aaronson strong-simulation [2];

$$\lambda \equiv \begin{array}{c} \{g_i\}_i \\ \\ \{S_i\}_i \end{array} \left| \begin{array}{cccc|cccc|c} a_{x_1,g_1} & a_{x_2,g_1} & \cdots & a_{x_n,g_1} & a_{z_1,g_1} & a_{z_2,g_1} & \cdots & a_{z_n,g_1} & \gamma(g_1) \\ a_{x_1,g_2} & a_{x_2,g_2} & \cdots & a_{x_n,g_2} & a_{z_1,g_2} & a_{z_2,g_2} & \cdots & a_{z_n,g_2} & \gamma(g_2) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{x_1,g_{n-1}} & a_{x_2,g_{n-1}} & \cdots & a_{x_n,g_{n-1}} & a_{z_1,g_{n-1}} & a_{z_2,g_{n-1}} & \cdots & a_{z_n,g_{n-1}} & \gamma(g_{n-1}) \\ \hline a_{x_1,S_1} & a_{x_2,S_1} & \cdots & a_{x_n,S_1} & a_{z_1,S_1} & a_{z_2,S_1} & \cdots & a_{z_n,S_1} & \gamma(S_1) \\ a_{x_1,S_2} & a_{x_2,S_2} & \cdots & a_{x_n,S_2} & a_{z_1,S_2} & a_{z_2,S_2} & \cdots & a_{z_n,S_2} & \gamma(S_2) \\ a_{x_1,S_3} & a_{x_2,S_3} & \cdots & a_{x_n,S_3} & a_{z_1,S_3} & a_{z_2,S_3} & \cdots & a_{z_n,S_3} & \gamma(S_3) \end{array} \right. \tag{5.17}$$

Where the binary-sympletic representation has been used to encode Pauli operators. From this we can see that an ontic state can be stored in $(2n + 1)(n + 2)$ bits, which satisfies the bound proven by Karanjai *et al.* [40]. The ontology is symmetric as each ontic state is in the support of exactly 3 quantum states, given by the stabilizer groups $\mathcal{S}(\psi_1)$, $\mathcal{S}(\psi_2)$, and $\mathcal{S}(\psi_3)$.

This construction drastically reduces the size of the ontology of the model. It also provides a route to investigating whether such a model can be used as the basis for a weak-simulation scheme, which will be covered in a follow up paper.

5.10   Generalized Contextuality

Generalized contextuality is an extension of traditional contextuality to preparations, transformations, and probabilistic ontological models. This generalization is defined via the concept of operational equivalences. We say that two (experimental) physical operations are *operationally equivalent*, denoted $\cong$, if and only if they give the same experimental statistic, regardless of the choice of experimental procedure:

Preparations:	Two preparation procedures P and P' are operationally equivalent, $(P \cong P')$, *iff* $\Pr(k	P, T, M) = \Pr(k	P', T, M)$, $\forall T, M$.
Transformations:	Two transformation procedures T and T' are operationally equivalent, $(T \cong T')$, *iff* $\Pr(k	P, T, M) = \Pr(k	P, T', M)$, $\forall P, M$.
Measurements:	The outcome k of two measurement procedures $k \in M$ and $k \in M'$ is operationally equivalent, $([k, M] \cong [k, M'])$, *iff* $\Pr(k	P, T, M) = \Pr(k	P, T, M')$, $\forall P, T$.

The assumption of generalized non-contextuality then states that any two operationally equivalent procedures in our physical theory should be represented by the same object in an ontological model. Therefore we say an ontological model is *preparation non-contextual* (PNC) if;

$$\mu_P = \mu_{P'} \Leftrightarrow P \cong P'. \tag{5.18}$$

Similarly it is *transformation non-contextual* (TNC) if;

$$\Gamma_T = \Gamma_{T'} \Leftrightarrow T \cong T'. \tag{5.19}$$

And *measurement non-contextual* (MNC) if;

$$\xi_{k,M} = \xi_{k,M'} \Leftrightarrow [k, M] \cong [k, M']. \tag{5.20}$$

As previously mentioned the n-qubit stabilizer formalism exhibits all forms of contextuality. Therefore it is not surprising that the model presented in this paper does not satisfy *any* of these requirements. Hence it is a preparation-transformation-measurement contextual ontological model.

5.10.1 The model and generalized contextuality

Preparation Contextuality

The model is preparation contextual for $n > 1$. To see this we use what we call *Pauli eigenbases*. A Pauli eigenbasis $\{\psi_i\}_i$ is a set of stabilizer states such that we can write the stabilizer groups of all eigenstates as $\mathcal{S}(\psi_i) = \langle \pm G_1, \pm G_2, ..., \pm G_n \rangle$, where $\{G_i \in \widetilde{\mathcal{P}}_n\}$ are the same for all eigenstates. We call these Pauli eigenbases as they are exactly the joint eigenstates of sets of commuting Pauli operators. However it should be noted that these are not the only orthonormal bases we can construct in the stabilizer formalism. The

PBR POVM elements give a clean example of an orthonormal basis that does not have the structure of a Pauli eigenbasis[9].

So without loss of generality let us consider the two Pauli eigenbases;

$$\{\psi_{Z,j}\} \cong \left\{ \mathcal{S}(\psi_{Z,j}) = \left\langle (-1)^{f(j,i)} Z_i \right\rangle_{i=1,..,n} \right\},$$

$$\{\psi_{X,k}\} \cong \left\{ \mathcal{S}(\psi_{X,k}) = \left\langle (-1)^{f(k,i)} X_i \right\rangle_{i=1,..,n} \right\},$$

where $Z_i = \mathbb{I}_{j \neq i} \otimes Z_i$ and similarly for X_i, and $(-1)^{f(i,j)}$ is a function that maps index j to all possible ± 1 phases on each generator. These are the Z (computational) and X eigenbases respectively and clearly share no non-identity stabilizer element, i.e. $\nexists S, j, k$ such that $S \in \mathcal{S}(\psi_{Z,j}) \backslash \mathbb{I}$ and $S \in \mathcal{S}(\psi_{X,k}) \backslash \mathbb{I}$. Therefore by noting the definition of the support of a stabilizer state in the model we can infer that $\mu_{\psi_{Z,j}} \cap \mu_{\psi_{X,k}} = \emptyset$, $\forall j, k$.

To demonstrate that the model is preparation contextual consider that both eigenbases can be used to prepare the maximally mixed state;

$$\frac{1}{2^n} \mathbb{I} = \sum_j \frac{1}{2^n} |\psi_{Z,j}\rangle \langle \psi_{Z,j}| = \sum_k \frac{1}{2^n} |\psi_{X,k}\rangle \langle \psi_{X,k}| = \frac{1}{2^n} \mathbb{I}.$$

Therefore the representation of the maximally mixed state depends on which eigenbasis we perform the preparation in;

$$\mu_{\mathbb{I}/2^n}^{(Z)} = \sum_j \frac{1}{2^n} \mu_{\psi_{Z,j}} \neq \sum_k \frac{1}{2^n} \mu_{\psi_{X,k}} = \mu_{\mathbb{I}/2^n}^{(X)},$$

by the disjointness of the pure state preparations. Therefore the model is preparation contextual. Strangely, this implies $\text{supp}(\mu_{\mathbb{I}/2^n}^{(Z)}) \cap \text{supp}(\mu_{\mathbb{I}/2^n}^{(X)}) = \emptyset$. However, we can define a *canonical* representation of the maximally mixed state via a uniform distribution over all possible representations.

The above analysis does not apply in the $n = 1$ case as the stabilizer part of the ontology is trivial, i.e. for a single qubit $\Lambda_{n=1} = \{(\mathbb{I}, \gamma) | \gamma \in \gamma_{n=1}\}$ and $\mathbb{I} \in \mathcal{S}(\rho)$, $\forall \rho \in \mathcal{S}(\mathcal{H}_2)$. Therefore we can effective discard it, making the $n = 1$ case preparation non-contextual as all bases span the space of phase functions, note this spanning nature of a basis also holds for $n > 1$.

---

[9]To construct these bases you leverage the fact to stabilizer states ρ_1 and ρ_2 are orthogonal *iff* $P \in \mathcal{S}(\rho_1)$ and $-P \in \mathcal{S}(\rho_2)$.

Transformation Contextuality

The model inherits the transformation contextuality of the single qubit stabilizer subtheory [51]. This implies, via embedding, the model is transformation contextual for all n.

Recently there has been proposals to strengthen the definition of transformation contextuality to be more akin to traditional contextuality [53]. This definition would class an operational theory as transformation contextual if the representation of a unitary transformation was dependent on the set of unitaries it was performed with. Under this definition the model is transformation non-contextual.

Measurement Contextuality

The model up to this point has been constructed such that only rank 2^{n-1} measurements are permissible, i.e. binary outcome measurements. This restriction has been imposed on the model as it directly corresponds to the measurements allowed within the stabilizer formalism. However any model of binary outcome measurements are trivially measurement non-contextual, as there is only one context a projector can be measured in.

Considering, a sequence of rank-2^{n-1} PVMs can be used to construct rank-1 PVMs. For example we can construct a computational basis measurement for two-qubits by measuring $Z\mathbb{I}$, with outcome k_1, followed by $\mathbb{I}Z$, with outcome k_2, which will project any state onto $|k_1 k_2\rangle$ with the correct probability. Alternatively we could have measured $\mathbb{Z}I$ followed by ZZ, with outcome k_3, which will project onto the state $|k_1(k_3 \oplus k_1)\rangle$. This gives two contexts in which the computational basis can be measured.

By using the operational elements of the model we can construct *effective* response functions that correspond to measuring any rank 2^k, $k \in \{1, ..., n-1\}$ observable with stabilizer eigenspaces. Here we will present this construction for two sequential measurements. However, it can easily be extended to longer measurement sequences. We consider a sequence of two measurements $M_1 \rightarrow M_2$ such that $[M_1, M_2] = 0$, this implies that both are diagonal in some basis. Letting the outcome of each measurement be k_1 and k_2,

respectively, we can construct effective response function for the joint measurement via;

$$\xi_{(M_1 \to M_2)}(k_1, k_2 | \lambda) = \Pr(k_1, k_2 | M_1, M_2, \lambda) = \Pr(k_2 | k_1, M_1, M_2, \lambda) \Pr(k_1 | M_1, \lambda),$$

$$= \int_\Lambda d\lambda' \Pr(k_2, \lambda' | k_1, M_1, M_2, \lambda) \Pr(k_1 | M_1, \lambda),$$

$$= \int_\Lambda d\lambda' \Pr(k_2 | M_2, \lambda') \Pr(\lambda' | k_1, M_1, \lambda) \Pr(k_1 | M_1, \lambda),$$

$$= \int_\Lambda d\lambda' \xi_{M_2}(k_2 | \lambda') \Gamma_{k_1 | M_1}(\lambda', \lambda) \xi_{M_1}(k_1 | \lambda), \tag{5.21}$$

where $\xi_{M_i}(k_i | \lambda)$ are the response functions for a binary outcome measurements and $\Gamma_{k|M}$ is the post-measurement transition map. Therefore the effective response functions are dependent on the sequence of measurements and choice of rank-2^{n-1} measurements used. Implying the model is measurement contextual.

Alternatively, to see how the model must be measurement contextual consider that traditional non-contextuality is the conjunction of measurement non-contextuality and outcome determinism. Therefore as the model is outcome deterministic and traditionally contextual, it must be measurement contextual[10].

---

[10]A final possible way to derive measurement contextuality is to note a projector can be associated to an outcome in many different measurement procedures. For example, $\Pi_{00} = |00\rangle \langle 00|$ is an a possible outcome of measuring each qubit in the computational basis, e.g. $M_1 = Z\mathbb{I}$ and $M_2 = \mathbb{I}Z$, and it is an outcome of a PBR-style measurement, e.g. $M_1 = ZZ$ then adaptively measure M_2 according to; $\mathcal{O}(M_1) = 0 \longrightarrow M_2 = Z\mathbb{I}, \mathcal{O}(M_1) = 1 \longrightarrow M_2 = XX$.

Chapter 6

Is the Stabilizer Subtheory the Limit of ψ-Epistemic Models?

6.1 Chapter Preamble

In this chapter we present a result that was derived during the construction of the contextual ψ-epistemic model of the n-qubit stabilizer formalism. Namely that for any ontological model of the n-qubit stabilizer formalism there must exist some region of the ontology that encodes $n-1$ generators of a stabilizer group, and that any measurement will "drive" the ontological representation of a state to this region. There is some irony in this work, as the ψ-epistemic model, presented in the previous chapter, started life in a version of this proof, which I believed was wrong, but turned out to be correct.

The work presented here was solely completed by myself, with supervision from Joseph Emerson. It will appear in a future publication.

Is the stabilizer subtheory the limit of ψ-epistemic models?

Piers Lillystone[1], and Joseph Emerson[2,3]

[1]*Institute for Quantum Computing and Department of Physics and Astronomy,*
University of Waterloo, Waterloo, Ontario N2L 3G1, Canada
[2]*Institute for Quantum Computing and Department of Applied Mathematics,*
University of Waterloo, Waterloo, Ontario N2L 3G1, Canada
[3] *Canadian Institute for Advanced Research, Toronto, Ontario M5G 1Z8, Canada*

Quantum computation has the potential to offer exponential computational advantage over classical computation. However, the physical source of this advantage is far from well understood, despite many proposals. In this letter we show that the efficiently simulable n-qubit stabilizer subtheory is on the verge of being ψ-ontic. However, it is know to admit a ψ-epistemic ontological model. Combining this result with the recent result demonstrating that previously constructed ψ-epistemic models of full quantum theory cannot describe state-update after measurement leads us to a tantalizing question: Is the impossibility of a ψ-epistemic model the source of a quantum computer's power?

6.2   Introduction

Understanding the resources behind the exponential quantum computational speed-up is more critical than ever, especially with intermediate-scale quantum computing on the verge of being realized in an experimental setting [65, 59]. However, there is little consensus on the source of this speed up. What is better understood is the subtheories of quantum theory that offer no computational advantage over classical computation. Namely the efficiently classically simulable stabilizer subtheory [31], which underpins most quantum error-correction techniques [60], and recently has been used as the basis for simulating universal quantum computation via stabilizer rank [14, 13, 67, 10].

Contextuality [7, 43, 55, 63, 56, 75], a concept originating from foundations of quantum theory research, has emerged as one of the more intriguing possible explanations for the exponential speed-up of quantum computation. Contextuality is necessary to acheive speed-up by injection of magic states into Clifford circuits [39, 25, 9, 71]; and quantifies the computational advantage for both magic-state and measurement based models of

quantum computation [4, 70, 82, 68, 83, 9, 61, 3, 21]. However, contextuality cannot be a sufficient resource as the qubit stabilizer subtheory offers a quantum computer no exponential computational advantage over a classical counterpart [31, 2], yet is contextual under all definitions of contextuality [55, 63, 56, 4, 38, 45, 51].

This prompts us to consider alternative possibilities for the source on quantum computational speed-up. In this letter we apply another concept from the foundations of quantum theory to the question of what provides the power behind exponential quantum computational speed-up: The distinction between ψ-ontic and ψ-epistemic ontological models [35, 36, 50]. In broad strokes, in a ψ-ontic ontological model th physical state of system always uniquely identifies which quantum state the system is in. Conversely in a ψ-epistemic model the physical state may not uniquely define the quantum state of the system. ψ-epistemic models are often considered to give classical accounts of quantum theory. While no proof[1] of the requirement of a ψ-ontic ontological model of quantum theory is known, there is a growing body of evidence that quantum theory requires a ψ-ontic ontological model [66, 1, 34, 23, 72]. Therefore, the distinction between ψ-ontic and ψ-epistemic models offers a fertile research avenue for investigating the source of the exponential speed-up offered by quantum computation.

Recently we demonstrated that all previously constructed ψ-epistemic models of full quantum theory fail when accounting for state-update after measurement [72]. We also constructed a ψ-epistemic model of the n-qubit stabilizer subtheory which evades such proofs [52]. Here we extend these result and prove that any ontological model of the n-qubit stabilizer subtheory requires at least $n-1$ generators to be encoded in an ontological model's ontology, to correctly reproduce the experimental statistics of the subtheory. Considering that n generators are required to uniquely specify a stabilizer state and that adding any unitary outside the Clifford group can promote the subtheory to universality [57, 58] the requirement of encoding $n-1$ generators in the ontology places the n-qubit stabilizer subtheory at the verge of ψ-onticity.

Additionally the result applies to weak simulation that never output incorrect answers, i.e. if $\text{Tr}(\Pi_{k,M}\rho) \in \{0,1\}$ then $\text{Pr}_{\text{sim}}(k|M,\rho) \in \{0,1\}$ for all possible simulation rounds [81, 40]. This implies that a weak simulation of the n-qubit stabilizer subtheory must also eventually encode $n-1$ generators in its internal memory, which may have implications for the efficiency of the recently developed stabilizer rank simulations. Lastly, we improve the lower bound on the size of the ontology required by an ontological model of the n-qubit stabilizer subtheory given by Karanjai *et al.* [40].

---

[1]That doesn't rely on additional assumptions.

6.3 Background Material

Ontological models formalism is defined under the mathematical framework of *operational theories*. An operational theory mathematically describes the outcome statistics of sets of experiments, which are decomposed into sequences of preparations, transformations, and measurements. That is, they give the probability of observing outcome k of a measurement M, conditioned on the preparation P, and transformation T performed in the experiment, i.e. the conditional probability $\Pr(k|P, T, M)$. For example, in quantum theory an experiment where we prepare a density matrix ρ, apply a completely-positive trace-preserving (CPTP) map Φ, and measure a positive-operator-valued measure (POVM) $M = \{E_k\}$ has operational statistics given by the Born rule $\Pr(k|\rho, \Phi, M) = \mathrm{Tr}(E_k \Phi(\rho))$.

The most basic assumption an ontological model (OM) makes is that there exists a set of physical, *ontic*, states $\Lambda = \{\lambda\}$ that fully describe the physical properties of a system. In an OM a preparation procedure, P, probabilistically prepares a physical state consistent with P. Therefore we can represent the preparation procedure P by the probability density function $\mu_P(\lambda) = \Pr(\lambda|P)$, where $\int_\Lambda \mu_P(\lambda)d\lambda = 1$. We define the support of a preparation to be $\mathrm{supp}(\mu_P) = \{\lambda|\, \mu_P(\lambda) > 0\}$. A transformation T stochastically maps ontic states to ontic states. We represent T by a transition matrix $\Gamma_T(\lambda', \lambda) = \Pr(\lambda'|\lambda, T)$, where $\int_\Lambda \Gamma_T(\lambda', \lambda)d\lambda' = 1, \forall \lambda$. A measurement M in an OM outputs an outcome k probabilistically, conditioned on the ontic state of the system. Therefore M is represented by a set of response functions $\{\xi_k^M(\lambda) = \Pr(k|\lambda, M)\}$, where $\sum_k \xi_k^M(\lambda) = 1, \forall \lambda$.

We say an OM successfully reproduces the outcome statistics of an operational theory if;

$$\int_\Lambda \xi_k^M(\lambda)\Gamma_T(\lambda', \lambda)\mu_P(\lambda)d\lambda d\lambda' = \Pr(k|P, T, M), \tag{6.1}$$

for all preparations, transformations and measurements in the operationally theory, $\forall P \in \mathcal{P}, \forall T \in \mathcal{T}$, and $\forall M \in \mathcal{M}$ respectively. Additional care must be taken when representing measurement update maps, because they are only defined for ontic states that can give the corresponding outcome. However, the results presented here do not rely on these finer details and a full review of their implications can be found in [72].

In this paper we are interested in the operational theory defined by the n-qubit stabilizer subtheory of quantum theory [60], i.e. $\Pr(k|\rho, C, M)$ where $\rho \in \mathcal{S}(\mathcal{H}_{2^n})$, $C \in Cl_{2^n}$, and $M \in \widetilde{\mathcal{P}}_n$.

The n-qubit stabilizer subtheory can be defined by starting with the projective Pauli

group;

$$\widetilde{\mathcal{P}}_n = \{\otimes_i^n P_i | P_i \in \{\mathbb{I}, X, Y, Z\}\} = \mathcal{P}_n \backslash U(1). \tag{6.2}$$

From this we define the set stabilizer operators to be $\mathcal{S}_n = \left\{\pm P \middle| P \in \widetilde{\mathcal{P}}_n\right\} \backslash -\mathbb{I}$, where $-\mathbb{I}$ is excluded as it stabilizes no state.

The set of pure states in the subtheory, $\mathcal{S}(\mathcal{H}_{2^n})$, are the joint $+1$ eigenstates of the stabilizer operators. As each eigen-subspace of the stabilizer operators has dimension 2^{n-1}, n commuting stabilizer operators, $g_i \in \mathcal{S}_n$, are required to uniquely specify a pure state;

$$|\psi\rangle\langle\psi| = \frac{1}{2^n} \prod_{i=1}^n (\mathbb{I} + g_i) = \frac{1}{2^n} \sum_{S \in \mathcal{S}(\psi)} S,$$

where $\mathcal{S}(\psi)$ is the Abelian group defined by the n stabilizer operators, a particular choice of generators for the group.

Permissible transformations in the subtheory are defined by the normalizer of the stabilizer operators, the Clifford group;

$$Cl_{2^n} = \left\{U \in U(2^n) \middle| USU^\dagger \in \mathcal{S}_n, \forall S \in \mathcal{S}_n\right\}.$$

The set of measurement operators is exactly the projective Pauli group, $\widetilde{\mathcal{P}}_n$.

6.4 State-upate in Ontological Models of the n-Qubit Stabilizer Subtheory

We start by considering an arbitrary ontological model of the n-qubit stabilizer formalism that crucially can model measurement update correctly.

From these very simple assumptions we immediately infer the following;

Lemma 6.4.1 *Given a pure stabilizer state ψ with stabilizer group;*

$$\mathcal{S}(\psi) = \{(-1)^{\gamma(b)} P_b | [P_a, P_b] = 0, P_b \in \widetilde{\mathcal{P}}_n,$$
$$\gamma : \widetilde{\mathcal{P}}_n \to \mathbb{Z}_2\}, \tag{6.3}$$

it must be the case that $\forall \lambda \in supp(\mu_\psi)$ the response functions for Pauli observables are

> *constrained as;*
>
> $$\xi_{\gamma(b), P_b}(\lambda) = \begin{cases} 1, & \forall P_b \in \mathcal{S}(\psi) \\ 0 & otherwise \end{cases}$$
>
> $$\implies v_\lambda(P_b) = \gamma(b). \tag{6.4}$$

Proof. Let $(-1)^{\gamma(b)} P_b \in \mathcal{S}(\psi)$ then using the trace-orthonormality of Pauli operators we have;

$$\mathrm{Tr}\left([\psi]\Pi_{\gamma(b), P_b}\right) = \frac{1}{2^{n+1}} \mathrm{Tr}\left(2\mathbb{I}\right),$$
$$= 1,$$
$$= \int_\Lambda \xi_{\gamma(b), P_b}(\lambda)\mu_{P_\psi}(\lambda)d\lambda,$$

where $[\psi] = |\psi\rangle\langle\psi| = \frac{1}{2^n}\sum_{a\in\mathcal{S}(\psi)}(-1)^{\gamma(a)}P_a$ and $\Pi_{\gamma(b), P_b} = \frac{1}{2}(\mathbb{I} + (-1)^{\gamma(b)}P_b)$ is the PVM that projects onto the $(-1)^{\gamma(b)}$ eigenspace of the P_b observable. However as the epistemic state is normalized over Λ;

$$\int_\Lambda \mu_{P_\psi}(\lambda)d\lambda = \int_{\mathrm{supp}(\mu_{P_\psi})} \mu_{P_\psi}(\lambda)d\lambda = 1.$$

As both μ_{P_ψ} and $\xi_{\gamma(a), P_a}$ are positive functions over Λ it must be the case that $\xi_{\gamma(a), P_a}(\lambda) = 1, \forall\lambda \in \mathrm{supp}(\mu_\psi)$. ∎

As the response functions are outcome deterministic on the support of ψ and Pauli observables are binary outcome measurements, we can define a value assignment as given by eq. (6.4). It should be noted this value assignment is *not* a value assignment in the sense of KS non-contextuality because it is only defined over elements of the group and not the full projective Pauli group. Further, we have used a binary representation of the value assignment, so $0 \leftrightarrow +1$ and $1 \leftrightarrow -1$ and addition is performed mod 2.

In the stabilizer formalism the measurement update rules of measuring a Pauli observable $M \in \widetilde{\mathcal{P}}_n$ can be expressed as follows. Given any pure stabilizer group $\mathcal{S}(\psi)$ and a Pauli operator M it is always possible to express the group as $\mathcal{S}(\psi) = \langle G, h\rangle$ where $[G_i, M] = 0$, $\forall G_i \in G$ and G is an Abelian subgroup of $\mathcal{S}(\psi)$ with $\pm M \notin G$. Therefore the updated stabilizer group post measurement of M with outcome k is $\mathcal{S}(\psi') = \langle G, (-1)^k M\rangle$, where k is the outcome of the measurement. I.e. we remove any non-commuting operators in the

pre-measurement group and add $(-1)^k M$, and compositions of M with G, into the group in their place.

We can express the post-measurement group as;

$$\mathcal{S}(\psi') = \Big\{ \Big\{ (-1)^{\gamma(g)} P_g, (-1)^{\gamma(g)+k+\beta(g,m)} P_{g+m} \Big\}$$
$$\big| [P_g, M] = 0, P_g \in \widetilde{G} \Big\},$$

where we have used the definition $P_a P_b = (-1)^{\beta(a,b)} P_{a+b} | [P_a, P_b] = 0$ to perform the group multiplication and $\widetilde{G} = G \backslash U(1)$.

We can apply a similar logic as we did to the response functions for the pre-measurement state. Going straight to value assignments we have $\forall \lambda' \in \mathrm{supp}(\mu_{\psi'})$;

$$v_{\lambda'}(P_g) = \gamma(g) = v_\lambda(P_g), \qquad\qquad \forall P_g \in \widetilde{G}$$
$$v_{\lambda'}(P_g M) = \gamma(g) + k + \beta(g, m), \qquad\qquad \forall P_g \in \widetilde{G}$$

The post-measurement update map must perform the transformation $\Gamma_{k|M} : \mathrm{supp}(\mu_\psi) \to \mathrm{supp}(\mu_{\psi'})$, and therefore must functionally satisfy the value assignment update as given above, i.e. $\Gamma_{k|M,\psi} : v_{\lambda \in \mathrm{supp}(\mu_\psi)} \to v_{\lambda' \in \mathrm{supp}(\mu'_\psi)}$, $\forall \lambda \in \mathrm{supp}(\mu_\psi)$.

To leverage these insights we use the following fact:

Fact 6.4.1 *In any ontological model that reproduces the statistics of the stabilizer formalism the ontic states can encode information about group elements of the stabilizer group.*

I.e. we can define a surjective function from the ontology to the set of Abelian subgroups of the Pauli group, $\mathcal{A}(\mathcal{P}_n)$;

$$f : \Lambda \to \mathcal{A}(\mathcal{P}_n). \tag{6.5}$$

For example, in a ψ-ontic model of the stabilizer formalism, i.e. the Gottesman-Knill theorem [31], this surjective function will map directly onto the maximal Abelian subgroups of the Pauli group. Conversely in the 8-state model [84, 11] the corresponding function maps every ontic state to the trivial group, as the stabilizer states with support on a given ontic state λ have non-commuting generators and hence λ encodes no information about the stabilizer elements. Utilizing this fact, rather than dealing directly with the ontic states, we can work with these subgroups rather than the ontic states.

Additionally, we can restrict our attention to the "hardest" case. By hardest case we mean the case in which a measurement update map would have the least information about the new elements of the post-measurement stabilizer group. Therefore, the update rule has the most difficulty in correctly assigning values to these new stabilizer elements.

> **Fact 6.4.2** *Knowledge of G or some some subgroup of G is the "hardest" case for a state update map.*

Consider the contrary, remembering that G is the group such that $\mathcal{S}(\psi) = \langle G, h \rangle$ where $\pm M \notin G$ and $[G, M] = 0$, if there are some elements of the known group G' that do not commute with M then we know by the stabilizer update rules that these elements are to be removed and replaced with $(-1)^k M$. Effectively allowing us to know the new elements of the stabilizer group, and hence their corresponding value assignments.

6.5 Any Ontological Model of the n-Qubit Stabilizer Subtheory Must Encode $n-1$ Generators in its Ontology

Firstly we use the assertion that any λ can only encode some subgroup of a maximal Abelian stabilizer group and the encoded group has $n - l$ generators, where l is an integer and $2 \leq l \leq n$, and that the group in question is the hardest case. The $l = 1$ case corresponds to our previously constructed ψ-epistemic model and therefore is possible. Let us label the subgroup λ encodes as G', this means that the the post-measurement stabilizer group can be expressed as $\mathcal{S}(\psi') = \left\langle G', G'^{\perp}, (-1)^k M \right\rangle$, where $G'^{\perp}$ is the unique group that extends $\langle G', M \rangle$ to $\mathcal{S}(\psi')$, i.e. $[g'^{\perp}, g] = 0$, $\forall g \in \langle G', M \rangle$ and G from the previous discussion on measurement update is given by $G = \left\langle G', G'^{\perp} \right\rangle$. However with only access to knowledge of G' and M there are many such group extensions;

$$\mathcal{G}_M^{\perp} = \{H \mid [h, M] = [h, g'] = 0; \, \forall h \in H, \, \forall g' \in G'\}.$$

For example for 2-qubits if $G' = \{\mathbb{I}\}$ and $M = YY$ then:

$$\mathcal{G}_{YY}^{\perp} = \{\langle \pm \mathbb{I}Y \rangle, \langle \pm Y\mathbb{I} \rangle, \langle \pm XZ \rangle,$$
$$\langle \pm ZX \rangle, \langle \pm XX \rangle, \langle \pm ZZ \rangle\}.$$

It general it may be possible that $\exists g \in G'^\perp$ and another group extension $G_i{}^\perp \in \mathcal{G}_M^\perp$ such that $g + M = g^{(i)} \in G_i{}^\perp$. For example in the possible group extensions given above $XX.YY = -ZZ$. By the constraints on the update rule in the ontological model given previously, without knowledge of the group G we must perform the update rule for all possible groups in $\mathcal{G}_M^\perp$, as any extension could be the correct extension. This implies, given a measurement of M with outcome k, we must update as follows;

$$v_{\lambda'}(g + M) = v_\lambda(g) + k + \beta(g, M),$$
$$v_{\lambda'}(g^{(i)} + M) = v_\lambda(g^{(i)}) + k + \beta(g^{(i)}, M),$$

where this update must be performed for all $g \in G'^\perp$ and all $g^{(i)} \in G_i{}^\perp$. However as $g \in G$ by assumption, it must be the case that $v_\lambda(g) = v_{\lambda'}(g) = v_{\lambda'}(g^{(i)} + M) = v_\lambda(g^{(i)}) + k + \beta(g^{(i)}, M)$ and as $v_\lambda(g)$ is a definite value this implies the value assignment on $v_\lambda(g^{(i)})$ must be fixed prior to measurement too. As we are able to choose M freely this implies that for a stabilizer state with stabilizer group $\mathcal{S}(\psi)$ we must have a fixed value assignment for all $\lambda \in \mathrm{supp}(\mu_\psi)$ that satisfies;

$$v_\lambda(P_a + P_b) = v_\lambda(P_a) + v_\lambda(P_b) + \beta(P_a, P_b),$$
$$\forall P_b \in \mathcal{S}(\psi), \forall P_a \in \widetilde{\mathcal{P}_n}, \, | \, [P_a, P_b] = 0. \tag{6.6}$$

Re-expressing this as a constraint on the value assignments for a state we have $v_\lambda(P_a + P_b) + v_\lambda(P_a) + \gamma(b) = \beta(P_a, P_b)$. Note this value assignment above bares similarities to Kochen-Specker non-contextuality. However there are subtle differences: Firstly, an update rule can change the assignment after a measurement and therefore is not a non-contextual assignment. Secondly, this constraint does not constrain the value assignment for all possible sets of commuting Paulis, just those that commute with stabilizer group elements. For 2-qubits it can easily be verified that such an assignment is possible, one such assignment is used in the proof below.

The main result is structured as follows: First we show that at least one generator is required to be known for the 2 qubit case. Second, we show that this proof can be generalized to n qubits. Last, we apply the n-qubit theorem to derive a lower bound on the number of bits required to represent the ontology of any model of the n-qubit stabilizer subtheory.

> **Theorem 6.5.1** *In any ontological model of the 2-qubit stabilizer subtheory, the ontology must encode one generator of a stabilizer group.*
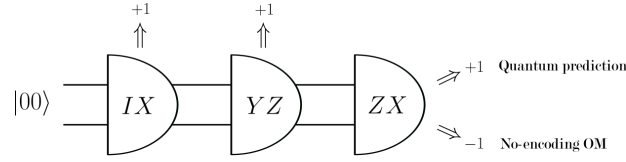
Figure 6.1: The circuit diagram of the two qubit measurement sequence used in the proof.

Proof. Without loss of generality[2], take the 2 qubit all zero state $|00\rangle$, which has stabilizer group $\mathcal{S}(|00\rangle) = \{\mathbb{I}, Z\mathbb{I}, \mathbb{I}Z, ZZ\}$. The set of measurements we will use to to prove the contradiction is $\mathbb{I}X$ followed by YZ then ZX. For these measurements we will condition on the $+1$ outcome for $\mathbb{I}X$ and YZ which implies the outcome of ZX is $+1$ with certainty, this can be verified by tracking stabilizer groups. However under the measurement update rules we will find that a model that cannot encode any generators in it's ontology must give the outcome -1 for the ZX measurement.

By assumption we can only encode $n-2$ generators for two qubits this implies the only group we can "know" is $G' = \langle \mathbb{I} \rangle$. Hence when performing the update rule we update on all Pauli operators that commute with the measurement, i.e.;

$$v_{\lambda'}(PM) = v_\lambda(P) + k + \beta(P, M),$$
$$\forall P \in \widetilde{\mathcal{P}}_2 \,|\, [P, M] = 0,$$

where k is the outcome of the measurement. Let us choose the value assignment $v_{|00\rangle}(P_a) = 0, \forall P_a \neq YY$ and $v_{|00\rangle}(YY) = 1$, which satisfies the constraint on the value assignment given by eq. (6.6) and must lie in the support of $\mu_{P_{|00>}}$ for the OM to give the right statistics.

For the first measurement the set of Pauli operators that commute with $\mathbb{I}X$ is;

$$\{Z\mathbb{I}, Y\mathbb{I}, X\mathbb{I}, \mathbb{I}, ZX, YX, XX, \mathbb{I}X\}. \tag{6.7}$$

As the β function is equal to 0 for all single qubit Pauli operators and by definition the value assignment is 0 on all these operators and $k = 0$, the value assignment doesn't change after the measurement of $\mathbb{I}X$.

Now we perform the measurement YZ, which will give a $+1$ outcome for the given value assignment. The set of operators that commute with YZ is;

$$\{\mathbb{I}, \mathbb{I}Z, Y\mathbb{I}, YZ, XX, ZY, ZX, XY\}, \tag{6.8}$$

---

[2]As a sequence of measurements used in the proof can be constructed for any 2-qubit input state.

where we are only interested in the last two. By the update rules we have;

$$v'(ZX) = v(XY) + k + \beta(XY, YZ) = 1,$$
$$v'(XY) = v(ZX) + k + \beta(ZX, YZ) = 1,$$

as $\beta(XY, YZ) = \beta(ZX, YZ) = 1$. Therefore a sequential measurement of ZX will give the outcome -1, which contradicts the quantum predictions. Note how to evade this condradiction the update rule would need to know which value assignment, $v(XY)$ or $v(ZX)$, should remain fixed, which effectively requires knowledge of one generator of the group.

∎

We note that theorem 6.5.1 does not explicitly rule out models which contain some set of ontic states that do not encode any generators. However, theorem 6.5.1 does imply that any such model will be "driven" to ontic states that do encode encode a generator, after any measurement, to avoid a possible contradiction. Therefore, under an ontological relaxation, i.e. where unnecessary ontology is removed, all models are constrained by theorem 6.5.1.

To extend this to the n qubit case we note the following: One possibility for the update rule is for it to perform some kind of sampling procedure. I.e. if a measurement M is performed and the group G is encoded in the pre-measurement ontic state λ, such that $[g, M] = 0, \forall g \in G$, then the sampling procedure samples from subgroups $G' \subset \langle G, M \rangle$, such that $|G'| \geq |G|^3$. However such a procedure will with non-zero probability violate the quantum predictions, via an embedding of theorem 6.5.1.

This implies that the ontology acts as a "memory" of previously performed measurements. However, even in a model where this is the case we still require $n-1$ generators to be encoded in the ontology.

> **Theorem 6.5.2** *In any ontological model of the n-qubit stabilizer subtheory, the ontology is required to encode $n-1$ generators of a stabilizer group.*

Proof. The general procedure of the proof will be to encode the 2-qubit proof in a sequence of measurements and show that if the ontology is acting as a memory then the effective stabilizer operator known for the 2-qubit subspace is $\mathbb{I}$. This is achieved by interleaving

---

[3]The assertion that $|G'| \geq |G|$ demands that the amount of information encoded by the ontic state can only increase or maintain the amount of information encoded by the ontology.
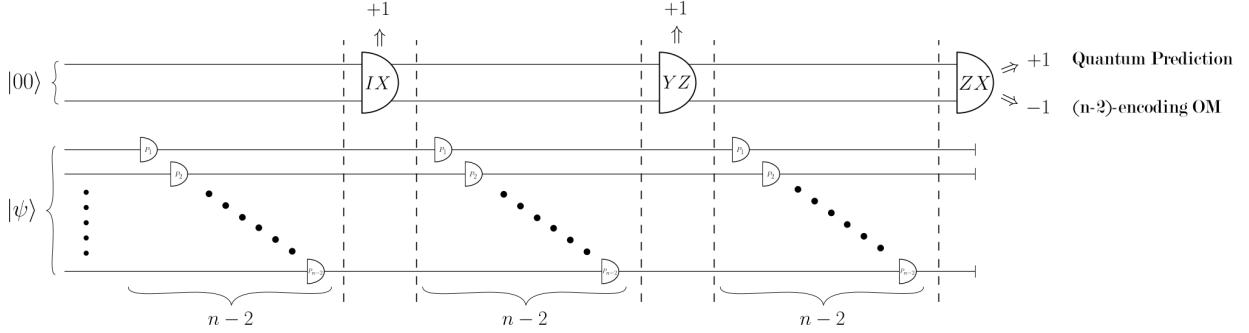
Figure 6.2: The circuit diagram of the n-qubit measurement sequence used in the proof.

each measurement in the 2-qubit proof with enough measurements on the other qubits to "purge" the 2-qubit measurement from the memory. Mathematically, if G is the stored group prior to any of the 2-qubit measurements then every element of G is of the form

$$g = P_{n-2} \otimes \mathbb{I}_2, \ \forall g \in G. \tag{6.9}$$

We let our initial state be $|\Psi\rangle = |\psi\rangle \otimes |00\rangle$, therefore the stabilizer group of this state is $S(\Psi) = \langle \mathcal{S}(\psi), \mathcal{S}(00)\rangle$. Again we assume w.l.g. that we can encode $n-2$ generators in the ontology. To demonstrate the purging process, let our initial encoded group be G_0 and $\exists P_{n-2} \otimes S \in G_0$, where $S \in \mathcal{S}(00)\backslash\mathbb{I}$. To purge these elements from the memory, but preserve the stabilizer group elements on the 2-qubit subspace, we need to perform a sequence of $n-2$ commuting measurements, M_i such that $[M_i, P_{n-2}\otimes S] = 0, \forall i$. A natural choice for this sequence is the set of single-qubit Pauli measurements $M_i = P_i \otimes \mathbb{I}_{n\neq i}$, where P_i is the i-th element of P_{n-2}. Therefore after performing such a sequence our new stabilizer group stored in memory is $G_{n-2} = \langle\{(-1)^{k_i}P_i\}_i\rangle$, where k_i indicates the outcome of M_i and we have suppressed the tensor notation. The stabilizer group of the state is $\mathcal{S}(\Psi') = \langle G_{n-2}, \mathcal{S}(00)\rangle$. This group therefore satisfies our requirement of "purging" the 2-qubit group from memory.

Recalling that the 2-qubit measurement sequence is $\mathbb{I}X \to YZ \to ZX$. After the first purging process we perform the first measurement in the sequence, $\mathbb{I}_{n-2}(\mathbb{I}X)$. As the ontology can encode $n-2$ commuting generators this measurement operator can be added to the memory. However we can perform a second purging sequence, as detailed above, between the $\mathbb{I}X$ and YZ measurements to ensure the stored group prior to the $\mathbb{I}_{n-2}(YZ)$ measurement satisfies eq. (6.9). This procedure can then be repeated prior to the third measurement in the sequence as well. Therefore at every stage of the measurement procedure the only information about the generators on the 2-qubit subspace stored in the

144

groups prior to each measurement is $\mathbb{I}$ and therefore theorem 6.5.1 can be embedded in a n-qubit measurement sequence. ∎

The results above have been presented in the context of ontological models. However, they equally apply to exact weak-simulations of the n-qubit stabilizer subtheory. Exact weak-simulations require that no output of the simulation contradicts the quantum predictions. Therefore a direct map between ontological models and exact weak-simulations can be drawn: The ontology, Λ, represents the sample space. Preparations, $\{\mu_P\}$, are the distributions from which samples are drawn. Stochastic maps, $\{\Gamma_T\}$, represent how a sample is evolved during simulation. Response functions, ξ_k^M, give the probability that the simulation outputs outcome k given sample λ.

Karanjai *et. al* gave a lower bound of $\frac{1}{2}n(n-1)$ for the number of bits required to specify the ontology for any ontological model of the n-qubit stabilizer formalism. theorem 6.5.2 allows us to improve this bound.

> **Corollary 6.5.3** *Any ontological model of the n-qubit stabilizer formalism requires at least $(2n+1)(n-1)$ bits to specify the ontology.*

Proof. A stabilizer operator is uniquely specified by $2n+1$ bits, and theorem 6.5.2 proves that at there must exist some region of the ontology which encodes all possible Abelian subgroups of the Stabilizer operators with $n-1$ generators. Therefore this set of ontic states requires at least $(2n+1)(n-1)$ bits to specify. ∎

6.6 Discussion

The stabilizer formalism is the basis upon which many quantum information tasks are designed, while being classically tractable in the computational sense. However, the addition of any non-Clifford unitary to the stabilizer formalism is enough to promote it to universal quantum computation. This makes the formalism ideal to study the transition between quantum and classical theories.

In this letter, we have shown that the stabilizer formalism is on the verge of being ψ-ontic, with $n-1$ generators required by any ontological model or exact weak simulation of the subtheory. This amount of encoded information is almost full information about the quantum state, with n generators required to uniquely specify the quantum state.

However, the n-qubit stabilizer formalism still admits a ψ-epistemic model [52]. Therefore, if additional elements are added to the subtheory it is hard to see how ψ-epistemicity could be preserved, as these additional element would require more information to be encoded in the ontology.

While the results presented here do not explicitly rule out a ψ-epistemic interpretation of full quantum theory. Recent results [72], showing that all previously constructed ψ-epistemic models fail to reproduce measurement update correctly, have reopened the question of whether a ψ-epistemic of full quantum theory is possible. This leads us to our final suggestion; The efficiently classically simulable stabilizer subtheories, for qubits and qudits, admit ψ-epistemic ontological models. However, the possibility of a ψ-epistemic ontological model of full quantum theory, which is assumed to be not efficiently simulable, seems increasingly distant. Therefore, does possibility/impossibility of ψ-epistemic/ψ-ontic model denote the boundary between quantum and classical computation? If so ψ-onticity would be *the* resource that enables quantum computational speed-up.

PART IV

Conclusion

In this thesis we have primarily investigated contextual phenomena that appear in seemingly classical systems. Demonstrating that contextuality is a far more prevalent feature of operational theories than previously thought. We have also explored the nature of the contextuality within these subtheories and provided routes beyond the current defintions of contextuality.

In part II we showed that the single-qubit stabilizer subtheory and single-rebit stabilizer subtheory both must admit a contextual ontological model, when transformations are included in the operational theory. Additionally, we showed that both of these subtheories also require a negatively represented quasi-probability representation, re-establishing the link between contextuality and negativity when transformations are included in a subtheory. Finally, we showed that, given a reasonable assumption that the identity transformation is represented by the identity matrix, almost-all unitaries in full quantum theory are negatively represented in any finite quasi-probability representation.

In part III we constructed a ψ-epistemic model of the n-qubit stabilizer subtheory and used insights from the model to show that in any ontological model of the n-qubit stabilizer subtheory at least $n - 1$ generators must be encoded in the ontology of the model. This result suggests that any subtheory of quantum theory that strictly contains the n-qubit stabilizer subtheory is more than likely to require a ψ-ontic ontological model, such that the additional informational content of the operational theory may be reproduced. Further evidence for this conjecture comes from a result published in collaboration with Joshua Ruebeck and Joseph Emerson [72], which reopens the door to a proof of the requirement of a ψ-ontic ontological model of quantum, without the requirement of additional assumptions.

Looking at the big picture of these results we can see that there is still much to learn about the boundary between classical theories and quantum theory, with a conclusive definition of what constitutes phenomena that are truly quantum still remaining elusive. However, with the inspiration provided by the field of quantum information, discovering this defining line is seeming closer than ever. If such a result can be found it would surely be of the same significance, and possibly of greater significance, as the celebrated Bell's theorem, while hopefully not sharing the same lag-period between discovery and acceptance as Bell's original theorem. It would also offer a tantalizing possibility, by understanding the fundamental implications of quantum theory on the nature of our physical reality, we may be able to construct a theory that unifies quantum theory and general relativity.

References

[1] Scott Aaronson, Adam Bouland, Lynn Chua, and George Lowther. ψ-epistemic theories: The role of symmetry. *Phys. Rev. A*, 88:032111, Sep 2013.

[2] Scott Aaronson and Daniel Gottesman. Improved simulation of stabilizer circuits. *Phys. Rev. A*, 70:052328, Nov 2004.

[3] Samson Abramsky, Rui Soares Barbosa, and Shane Mansfield. Contextual fraction as a measure of contextuality. *Phys. Rev. Lett.*, 119:050504, Aug 2017.

[4] Janet Anders and Dan E. Browne. Computational power of correlations. *Phys. Rev. Lett.*, 102:050502, Feb 2009.

[5] Nix Barnett and James P. Crutchfield. Computational mechanics of input–output processes: Structured transformations and the ϵ-transducer. *Journal of Statistical Physics*, 161(2):404–451, Oct 2015.

[6] Stephen D. Bartlett, Terry Rudolph, and Robert W. Spekkens. Reconstruction of gaussian quantum mechanics from liouville mechanics with an epistemic restriction. *Phys. Rev. A*, 86:012103, Jul 2012.

[7] John S. Bell. On the problem of hidden variables in quantum mechanics. *Rev. Mod. Phys.*, 38:447–452, Jul 1966.

[8] J.S. Bell. On the Einstein Podolsky Rosen Paradox. *Physics*, 1(3):195–200, 1964.

[9] Juan Bermejo-Vega, Nicolas Delfosse, Dan E. Browne, Cihan Okay, and Robert Raussendorf. Contextuality as a resource for models of quantum computation with qubits. *Phys. Rev. Lett.*, 119:120505, Sep 2017.

[10] Michael Beverland, Earl Campbell, Mark Howard, and Vadym Kliuchnikov. Lower bounds on the non-Clifford resources for quantum computations. 2019.

[11] Pawel Blasiak. Quantum cube: A toy model of a qubit. *Physics Letters A*, 377(12):847 – 850, 2013.

[12] Andreas Blass and Yuri Gurevich. Spekkens's Symmetric No-Go Theorem. 2015.

[13] Sergey Bravyi, Dan Browne, Padraic Calpin, Earl Campbell, David Gosset, and Mark Howard. Simulation of quantum circuits by low-rank stabilizer decompositions. 2018.

[14] Sergey Bravyi and David Gosset. Improved classical simulation of quantum circuits dominated by clifford gates. *Phys. Rev. Lett.*, 116:250501, Jun 2016.

[15] Sergey Bravyi and Alexei Kitaev. Universal quantum computation with ideal clifford gates and noisy ancillas. *Phys. Rev. A*, 71:022316, Feb 2005.

[16] Nicolas Brunner, Daniel Cavalcanti, Stefano Pironio, Valerio Scarani, and Stephanie Wehner. Bell nonlocality. *Rev. Mod. Phys.*, 86:419–478, Apr 2014.

[17] Adán Cabello. Experimentally testable state-independent quantum contextuality. *Phys. Rev. Lett.*, 101:210401, Nov 2008.

[18] Adán Cabello, Lars Eirik Danielsen, Antonio J. López-Tarrida, and José R. Portillo. Basic exclusivity graphs in quantum correlations. *Phys. Rev. A*, 88:032104, Sep 2013.

[19] Adán Cabello, Mile Gu, Otfried Gühne, and Zhen-Peng Xu. Optimal classical simulation of state-independent quantum contextuality. *Phys. Rev. Lett.*, 120:130401, Mar 2018.

[20] Adán Cabello, Simone Severini, and Andreas Winter. Graph-theoretic approach to quantum correlations. *Phys. Rev. Lett.*, 112:040401, Jan 2014.

[21] Lorenzo Catani and Dan E Browne. State-injection schemes of quantum computation in Spekkens ' toy theory. 2017.

[22] John F. Clauser, Michael A. Horne, Abner Shimony, and Richard A. Holt. Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.*, 23:880–884, Oct 1969.

[23] Roger Colbeck and Renato Renner. A system's wave function is uniquely determined by its underlying physical state. *New Journal of Physics*, 19(1):013016, jan 2017.

[24] Jeroen Dehaene and Bart De Moor. Clifford group, stabilizer states, and linear and quadratic operations over gf(2). *Phys. Rev. A*, 68:042318, Oct 2003.

[25] Nicolas Delfosse, Philippe Allard Guerin, Jacob Bian, and Robert Raussendorf. Wigner function negativity and contextuality in quantum computation on rebits. *Phys. Rev. X*, 5:021003, Apr 2015.

[26] A. Einstein, B. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.*, 47:777–780, May 1935.

[27] Christopher Ferrie. Quasi-probability representations of quantum theory with applications to quantum information science. *Reports on Progress in Physics*, 74(11):116001, 2011.

[28] Christopher Ferrie, Ryan Morris, and Joseph Emerson. Necessity of negativity in quantum theory. *Phys. Rev. A*, 82:044103, Oct 2010.

[29] Kathleen S. Gibbons, Matthew J. Hoffman, and William K. Wootters. Discrete phase space based on finite fields. *Phys. Rev. A*, 70:062101, Dec 2004.

[30] Daniel Gottesman. Stabilizer Codes and Quantum Error Correction. 1997.

[31] Daniel Gottesman. The Heisenberg Representation of Quantum Computers. *Proceedings of the XXII International Colloquium on Group Theoretical Methods in Physics*, 1:32–43, 1998.

[32] Daniel Gottesman. An Introduction to Quantum Error Correction and Fault-Tolerant Quantum Computation. 2009.

[33] Daniel M. Greenberger, Michael A. Horne, and Anton Zeilinger. Going beyond bell's theorem. *Bell's Theorem, Quantum Theory and Conceptions of the Universe*, pages 69–72, 1989.

[34] Lucien Hardy. Are quantum states real? *International Journal of Modern Physics B*, 27(01n03):1345012, jan 2013.

[35] Nicholas Harrigan and Terry Rudolph. Ontological models and the interpretation of contextuality. 2007.

[36] Nicholas Harrigan and Robert W. Spekkens. Einstein, incompleteness, and the epistemic view of quantum states. *Foundations of Physics*, 40(2):125–157, Feb 2010.

[37] Alexander S. Holevo. Bounds for the quantity of information transmitted by a quantum communication channel. *Problems of Information Transmission*, 9(3):177–183, 1973.

[38] Mark Howard, Eoin Brennan, and Jiri Vala. Quantum contextuality with stabilizer states. *Entropy*, 15(6):2340–2362, 2013.

[39] Mark Howard, Joel J. Wallman, Victor Veitch, and Joseph Emerson. Contextuality supplies the 'magic' for quantum computation. *Nature*, 509(7505):351, jun 2014.

[40] Angela Karanjai, Joel J. Wallman, and Stephen D. Bartlett. Contextuality bounds the efficiency of classical simulation of quantum processes. 2018.

[41] Matthias Kleinmann, Otfried Gühne, José R Portillo, Jan-Åke Larsson, and Adán Cabello. Memory cost of quantum contextuality. *New Journal of Physics*, 13(11):113011, 2011.

[42] Alexander A. Klyachko, M. Ali Can, Sinem Binicioğlu, and Alexander S. Shumovsky. Simple test for hidden variables in spin-1 systems. *Phys. Rev. Lett.*, 101:020403, Jul 2008.

[43] Simon Kochen and E. P. Specker. The problem of hidden variables in quantum mechanics. *Journal of Mathematics and Mechanics*, 17(1):59–87, 1967.

[44] Lucas Kocia and Peter Love. Discrete Wigner Formalism for Qubits and Non-Contextuality of Clifford Gates on Qubit Stabilizer States. 2017.

[45] Anirudh Krishna, Robert W Spekkens, and Elie Wolfe. Deriving robust noncontextuality inequalities from algebraic proofs of the Kochen-Specker theorem: the Peres-Mermin square. 2017.

[46] Ravi Kunjwal and Robert W. Spekkens. From the kochen-specker theorem to noncontextuality inequalities without assuming determinism. *Phys. Rev. Lett.*, 115:110403, Sep 2015.

[47] Jan-Åke Larsson. A contextual extension of spekkens' toy model. *AIP Conference Proceedings*, 1424(1):211–220, 2012.

[48] M. S. Leifer and O. J. E. Maroney. Maximally epistemic interpretations of the quantum state and contextuality. *Phys. Rev. Lett.*, 110:120401, Mar 2013.

[49] M. S. Leifer and R. W. Spekkens. Logical pre- and post-selection paradoxes, measurement-disturbance and contextuality. *International Journal of Theoretical Physics*, 44(11):1977–1987, Nov 2005.

[50] Matthew Saul Leifer. Is the Quantum State Real? An Extended Review of ψ-ontology Theorems. *Quanta*, 3(1):67, 2014.

[51] Piers Lillystone, Joel J. Wallman, and Joseph Emerson. Contextuality and The Single-Qubit Stabilizer Formalism. 2018.

[52] Piers Lillystone, Joel J. Wallman, and Joseph Emerson. Contextuality and the single-qubit stabilizer subtheory. *Phys. Rev. Lett.*, 122:140405, Apr 2019.

[53] Shane Mansfield and Elham Kashefi. Quantum advantage from sequential transformation contextuality. 2018.

[54] Michael D. Mazurek, Matthew F. Pusey, Kevin J. Resch, and Robert W. Spekkens. Experimentally bounding deviations from quantum theory in the landscape of generalized probabilistic theories. 2013.

[55] N. David Mermin. Simple unified form for the major no-hidden-variables theorems. *Physical Review Letters*, 65(27):3373–3376, 1990.

[56] N. David Mermin. Hidden variables and the two theorems of John Bell. *Reviews of Modern Physics*, 65(3):803–815, 1993.

[57] Gabriele Nebe, E. M. Rains, and N. J. A. Sloane. The invariants of the Clifford groups. 2000.

[58] Gabriele Nebe, Eric M. Rains, and Neil J. A. Sloane. *Self-Dual Codes and Invariant Theory*. Springer-Verlag Berlin Heidelberg, 1st edition, 2006.

[59] C. Neill, P. Roushan, K. Kechedzhi, S. Boixo, S. V. Isakov, V. Smelyanskiy, A. Megrant, B. Chiaro, A. Dunsworth, K. Arya, R. Barends, B. Burkett, Y. Chen, Z. Chen, A. Fowler, B. Foxen, M. Giustina, R. Graff, E. Jeffrey, T. Huang, J. Kelly, P. Klimov, E. Lucero, J. Mutus, M. Neeley, C. Quintana, D. Sank, A. Vainsencher, J. Wenner, T. C. White, H. Neven, and J. M. Martinis. A blueprint for demonstrating quantum supremacy with superconducting qubits. *Science*, 360(6385):195–199, 2018.

[60] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, New York, NY, USA, 10th edition, 2011.

[61] Cihan Okay, Sam Roberts, Stephen D Bartlett, and Robert Raussendorf. Topological proofs of contextuality in quantum mechanics. *Quantum Information and Computation*, 17:1135–1166, 2017.

[62] Hakop Pashayan, Stephen D. Bartlett, and David Gross. From estimation of quantum probabilities to simulation of quantum circuits. 2017.

[63] A Peres. Two simple proofs of the kochen-specker theorem. *Journal of Physics A: Mathematical and General*, 24(4):L175, 1991.

[64] John Preskill. Quantum computing and the entanglement frontier. 2012.

[65] John Preskill. Quantum Computing in the NISQ era and beyond. *Quantum*, 2:79, August 2018.

[66] Matthew F. Pusey, Jonathan Barrett, and Terry Rudolph. On the reality of the quantum state. *Nature Physics*, 8(475), 2012.

[67] Hammam Qassim, Joel J. Wallman, and Joseph Emerson. Clifford recompilation for faster classical simulation of quantum circuits. 2019.

[68] Robert Raussendorf. Contextuality in measurement-based quantum computation. *Phys. Rev. A*, 88:022322, Aug 2013.

[69] Robert Raussendorf, Juani Bermejo-Vega, Emily Tyhurst, Cihan Okay, and Michael Zurel. Phase space simulation method for quantum computation with magic states on qubits. 2019.

[70] Robert Raussendorf and Hans J. Briegel. A one-way quantum computer. *Phys. Rev. Lett.*, 86:5188–5191, May 2001.

[71] Robert Raussendorf, Dan E. Browne, Nicolas Delfosse, Cihan Okay, and Juan Bermejo-Vega. Contextuality and wigner-function negativity in qubit quantum computation. *Phys. Rev. A*, 95:052334, May 2017.

[72] Joshua B. Ruebeck, Piers Lillystone, and Joseph Emerson. Epistemic interpretations of quantum theory have a measurement problem. 2018.

[73] David Schmid, Robert W Spekkens, and Elie Wolfe. All the noncontextuality inequalities for arbitrary prepare-and-measure experiments with respect to any fixed sets of operational equivalences. 2017.

[74] David Schmid, Robert W. Spekkens, and Elie Wolfe. All the noncontextuality inequalities for arbitrary prepare-and-measure experiments with respect to any fixed set of operational equivalences. *Phys. Rev. A*, 97:062103, Jun 2018.

[75] Andrew W Simmons, Joel J Wallman, Hakop Pashayan, Stephen D Bartlett, and Terry Rudolph. Contextuality under weak assumptions. *New Journal of Physics*, 19(3):033030, 2017.

[76] Robert W. Spekkens. Contextuality for preparations, transformations, and unsharp measurements. *Physical Review A*, 71:052108, may 2005.

[77] Robert W. Spekkens. Evidence for the epistemic view of quantum states: A toy theory. *Phys. Rev. A*, 75:032110, Mar 2007.

[78] Robert W. Spekkens. Negativity and contextuality are equivalent notions of nonclassicality. *Phys. Rev. Lett.*, 101:020401, Jul 2008.

[79] Robert W. Spekkens. The status of determinism in proofs of the impossibility of a noncontextual model of quantum theory. *Foundations of Physics*, 44(11):1125–1155, Nov 2014.

[80] Jayne Thompson, Andrew J. P. Garner, Vlatko Vedral, and Mile Gu. Using quantum theory to simplify input–output processes. *npj Quantum Information*, 3(1), Feb 2017.

[81] Maarten Van Den Nest. Classical simulation of quantum computation, the gottesman-knill theorem, and slightly beyond. *Quantum Info. Comput.*, 10(3):258–271, mar 2010.

[82] Victor Veitch, Christopher Ferrie, David Gross, and Joseph Emerson. Negative quasi-probability as a resource for quantum computation. *New Journal of Physics*, 14(11):113011, 2012.

[83] Victor Veitch, S A Hamed Mousavian, Daniel Gottesman, and Joseph Emerson. The resource theory of stabilizer quantum computation. *New Journal of Physics*, 16(1):013009, 2014.

[84] Joel J. Wallman and Stephen D. Bartlett. Non-negative subtheories and quasiprobability representations of qubits. *Physical Review A*, 85(06):062121, jun 2012.

[85] Sixia Yu and C. H. Oh. State-independent proof of kochen-specker theorem with 13 rays. *Phys. Rev. Lett.*, 108:030402, Jan 2012.

APPENDICES

Appendix A

Mathematica Code: Finite-state Representation of the 2-Qubit Model

A.1 Code Overview

Reproduced here is an abridged version of the mathematica code that simulates the 2-qubit ψ-epistemic model. Much of the internal working of the code have been omitted in the interest of space and one of the examples has been included to show how the code functions. The full code base is Accessable on github, date last accessed July 24th 2019. As presented the comments for each section of the code are visible.

ψ-Epistemic Model of 2-Qubit Stabilizer Formalism: A Finite-State Machine Representation

Introduction

In this Mathematica script we provide a code to construct a graphical representation of the ψ-epistemic model of the 2-qubit stabilizer formalism. All core functional code is contained in the sections below. However, for ease of use a main setup script is included in the final section, which will initialise all the code and hide it (hint: for a new mathematica user run a cell by selecting it and pressing "shift+enter"). We have opted for this approach over initialization cells to improve compatibility.

The general functionality of the code is to allow a user to input a specified Clifford circuit (which can contain adaptive measurements and gates) and be given an output representing the model's evolution under the circuit. To demonstrate this we have provided 4 examples that capture most of the codes capabilities. The specification of the circuit input is given in the *circuit parser* section.

Finally functionality is provided to give a graphical representation of the 2-qubit model. The representation we have chosen is that of a finite-state machine. A finite-state machine can be thought of as a box with certain internal parameter settings, with any operation dependent only on the internal parameters of the machine. For example, a box with a series of lights representing whether an internal state is zero or one. A description of the graphical representation is given in the *graphical plotting functions* section.

Global Variable Definitions

This section contains all the global variable definitions. Such as color of the cells in the graphical representation.

SubHeader

Graphical Representation Variables

SubHeader

Generic Variables

Generic Functions

This section contains generic functions to map between different representations of Pauli operators.

The first is an index over the Pauli group (below) which is then converted to a 4 bit binary vector. This vector we refer to as the tuple representation, we use this as it is easy to switch between the tuple and index representation.

Index reference: 1<->II, 2<->IZ, 3<->IX, 4<->IY, 5<->ZI, 6<->ZZ, 7<->ZX, 8<->ZY, 9<->XI, 10<->XZ, 11<->XX, 12<->XY, 13<->YI, 14<->YZ, 15<->YX, 16<->YY

We also provide function here to switch between the tuple and index representations to the binary-sympletic representation (described in the stabilizer state preparation section).

SubHeader

Index-2-Tuple-2-Binary Symplectic Representation Maps

Graphical Plotting Functions

In this section we provide function to plot the representation of an ontic state. To do so we arrange a 4x4 grid with each axis representing a local Pauli operator, as follows;

II	IX	IZ	IY
XI	XX	XZ	XY
ZI	ZX	ZZ	ZY
YI	YX	YZ	YY

The colour of a cell represents whether the value assignment for a given Pauli operator is 0 or 1 (default is green=0, blue =1). The black dot in a cell indicates which Pauli Operator the Pauli component of the ontology is equal to.

There are two provided plotting functions;

FiniteStatePauli -> Plots the finite machine representation with Pauli operators indexing the axes.
FiniteStateRaw -> Plots the raw finite state machine representation.

Both take as arguments a list of lambdas, $\{\lambda\_0, \lambda\_1, \lambda\_2, ...\}$, and output a respectively ordered list of the

finite state machine's representation for each lambda.

Finite-State Machine Representation Plotting Functions

Stabilizer State Preparation

The below code deals with stabilizer states:
A generic input stabilizer state is defined by two generators,

$$G_1 = (x_{1,G_1}, x_{2,G_1} | z_{1,G_1}, z_{2,G_1} | p_{G_1}), \; G_2 = (x_{1,G_2}, x_{2,G_2} | z_{1,G_2}, z_{2,G_2} | p_{G_2})$$

which follow the usual binary symplectic definitions.

In code this is represented as a list with 3 elements;

```
gi = {{x_ 1 _gi, x_ 2 _gi}, {z_ 1 _gi, z_ 2 _gi}, p_gi};
 (*For example: -XZ*)
MinusXZ = {{1, 0}, {0, 1}, 1};
```

Stabilizer Group Functions

Functions to Sample λ Given a Stabilizer Group

Clifford Operations

Section implements Clifford evolution maps. Note: Some Clifford operation are implemented as sequences of previously defined Clifford operations.

Hard-coded Clifford Operations

The directly coded maps are as follows:

Single Qubit gates - Paulis, Hadamard, Phase
Two Qubit gates - CNOT1→2,CNOT2→1 (one function)

Note: GammaH, GammaS etc could all be rolled into a generic function, but separately is easier to address.

All maps take in an ontic state as argument and output the resultant ontic state. Additionally the Pauli map takes in a Pauli operator as an argument and the CNOT map takes in control arguments (first -

control, second - target).

Composite Clifford Operations

Below are implementations of Clifford operations that can be expressed as sequences of the hard-coded operations.

Measurement Outcome/Measurement Update Rules

In the below we provide the implementation of the measurement update rules.

The core functionality is the provided in the function *OnticMeasurementUpdate* which takes as an argument a Pauli observable and ontic state, it then outputs the updated ontic state.

Additionally functionality is provided to interrogate the measurement update graphically. See "Circuit Parser Example/Tutorial Code" in the examples section for a full explanation of the how this visualization is structured.

Circuit Parser

Overview:

The code in this section is designed parse an input stabilizer circuit and output a list of outcomes and a history of the ontic state.

The input is broken down into blocks - each block represents a preparation then choice of mmt or transformation (possibly adaptive);

0th input → Preparation: {List[Stabilizer groups],OPTIONAL:List[Weights to sample stabilizer groups]} - If no weights provided uniform randomly weights will be used.

Ex: Sample Computational basis with

P(00)=0.5,P(01)=0.2,P(10)=0.2,P(11)=0.1

→ {

{ {"IZ","ZI"},{"-IZ","ZI"},{"IZ","-ZI"},{"-IZ","-ZI"}

},

{0.5,0.2,0.2,0.1}

}

ith input → Choice of Transformation or Meausurement: {Choice("T","M"), Adaptive(True/False), [Adaptive=False→OperationToImplement(Clifford,Pauli Obs),

Adapative=True→ListofConditionalOperations]}

Operations string input: Paulis - "P_1P_2" (i.e. "XX","IZ"),

Cliffords - "IH", "HI", "IS", "SI", "CNOT12", "CNOT21", "CZ

OPTIONAL: random operations - weighting can be applied as in preparation case.

Apdative structure is as follows: { DefaultOperationList,

{{k_i's to condtion on},{outcomes to condition on},{Operation list}},

...}}

Note: Operation list can include weightings as usual.

For example: {"II", {

{{1,2}, {0,0}, "IH"},

{{3},{1}, "SI"}

}}

Will as a default apply the identity operation, otherwise if measurement 1 and 2's outcome were both 0 it will apply IH, and if measurement 2's outcome was 1 it'll apply SI.

Warning the code will not error if multiple conditionals are satisfied.

Functionality:

The two core circuit parser functions are *RunCircuit* and *RunCircuitFromInput,* both performing as expected.

RunCircuit takes as input a circuit of the form described above. The output is a 3 element list --- {{State Prepared, Operations implemented}, {List of raw outcomes from measurements}, {list of lambdas}}

RunCircuitFromInput is the same as above, except it additionally takes as input the output of a previous circuit and a circuit. Note the circuit input can contain the state preparation specification, however it will be ignored.

Input Parsers and Randomized Operation Selectors

Implement Operations

Run Circuit Functions

Main Setup and Examples

Code Setup (RUN THIS PRIOR TO ANY EXAMPLE)

Code to setup up simulator.

Note: all code sections that are evaluated are all automatically closed after evaluation. If you wish to edit/inspect the code you can find it in the appropriate section above.

```
In[1]:= ClearAll;
     (*Evaluate "Core" Cells*)
     nb = EvaluationNotebook[];
     NotebookFind[nb, "Core", All, CellTags];
     SelectionEvaluate[nb]

     (*Close all evaluated sections
       (NOTE this is applied at the level of sub(sub)sections)*)
     nbClose = SelectedNotebook[];
     NotebookFind[nbClose, "SubHeader", All, CellTags];
     FrontEndTokenExecute["SelectionCloseAllGroups"]
```

Examples

Each Example can be run independently once the code setup has been executed.

Circuit Parser Example/Tutorial Code

Below is an arbitrary implementation of all the circuit parsers features. This implementation can used as

a template for other circuits.

```
In[74]:= (*Example input to circuit parser*)
CircEx = {
    {{ {"IZ", "ZI"}, {"XX", "-YY"} }, {0.75, 0.25}},
    (*Prepare |00> or |B_00> with probability 0.75 or 0.25*)
        {"T", False, {{"HI", "SI"}}},
    (*Non-adaptive Transformation: Apply HI or SI with equal probability*)
        {"M", False, {{"IX"}}}, (*Non-adaptive Measurement of IX*)
        {"T", True, {{{"II"}}, {{{1}, {0}, {{"ZX"}}}}}},
    (*Apdative Transformation: If measurement outcome of IX
        (2nd circuit element but 1st measurement) is 0 apply ZX,
    otherwise apply Identity (Indicated by first operation)*)
        {"M", False, {{"XZ", "YY"}, {0.25, 0.75}}}, (*2nd Non-
     Apdative Measurement: note 4th element of circuit, but second measurement*)
        {"M", True, {{{"XX"}}, {{{1, 2}, {0, 0}, {{"IZ"}}}, {{1, 2}, {1, 1}, {{"ZX"}}}}}}}
     (*Apdative measurement; measure IZ if outcome 1 = outcome 2 = 0,
     measure ZX if outcome 1 = outcome 2 = 1, otherwise measure XX*)
    }; (*Note code is currently not built to handle conditioning on
   which random operation is applied, maybe later...*)

OutputCirc = RunCircuit[CircEx]
```

```
Out[75]= {{{IZ, ZI}, HI, IX, II, XZ, ZX}, {1, 1, 0},
    {{{0, 1, 0, 0}, {0, 0, 1, 0, 0, 0, 0, 0, 1, 1, 0, 0, 1, 0, 0, 1}},
     {{1, 0, 0, 0}, {0, 0, 1, 0, 1, 1, 0, 0, 0, 0, 0, 0, 1, 1, 0}},
     {{0, 0, 1, 0}, {0, 0, 1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 0, 1, 1, 0}},
     {{0, 0, 1, 0}, {0, 0, 1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 0, 1, 1, 0}},
     {{1, 0, 0, 1}, {0, 1, 1, 0, 0, 1, 0, 1, 0, 1, 0, 0, 1, 1, 1, 1}},
     {{0, 1, 1, 0}, {0, 1, 1, 0, 0, 1, 0, 1, 0, 1, 0, 0, 1, 1, 0, 1}}}}
```

OutputCirc contains all relevant information about the execution of the circuit.

The first element of the output contains the potential randomly chosen circuit operations (note: we do not distinguish between measurements and transformations here);

```
In[76]:= OutputCirc[[1]]
```

```
Out[76]= {{IZ, ZI}, HI, IX, II, XZ, ZX}
```

The second element contains the measurement outcomes (useful for gathering statistics);

```
In[77]:= OutputCirc[[2]]
```

```
Out[77]= {1, 1, 0}
```

The final element contains the sequence of ontic states the circuit mapped through;

```
In[78]:= OutputCirc[[3]]
```

```
Out[78]= {{{0, 1, 0, 0}, {0, 0, 1, 0, 0, 0, 0, 0, 1, 1, 0, 0, 1, 0, 0, 1}},
    {{1, 0, 0, 0}, {0, 0, 1, 0, 1, 1, 0, 0, 0, 0, 0, 0, 1, 1, 0}},
    {{0, 0, 1, 0}, {0, 0, 1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 0, 1, 1, 0}},
    {{0, 0, 1, 0}, {0, 0, 1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 0, 1, 1, 0}},
    {{1, 0, 0, 1}, {0, 1, 1, 0, 0, 1, 0, 1, 0, 1, 0, 0, 1, 1, 1, 1}},
    {{0, 1, 1, 0}, {0, 1, 1, 0, 0, 1, 0, 1, 0, 1, 0, 0, 1, 1, 0, 1}}}}
```

To visualize this sequence we call the FiniteStatePauli or FiniteStateRaw functions which will output a list of graphic representations (the finite-state machine representation) of the above ontic states.

In[79]:= **FiniteStatePauli[OutputCirc⟦3⟧]**
FiniteStateRaw[OutputCirc⟦3⟧]

Out[79]=

Out[80]=

165
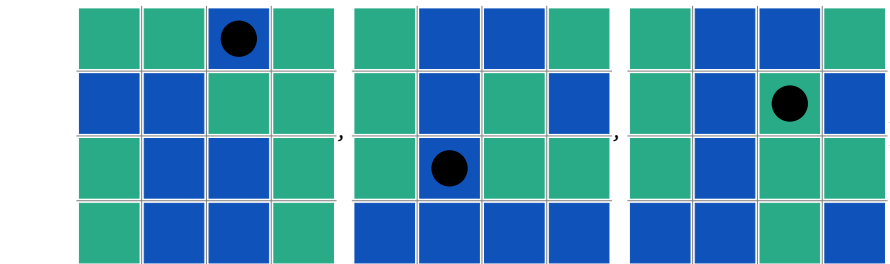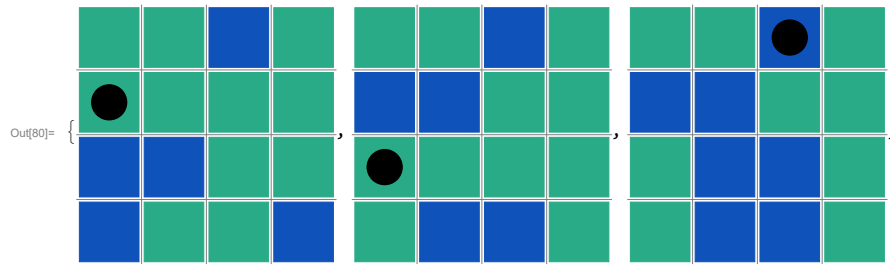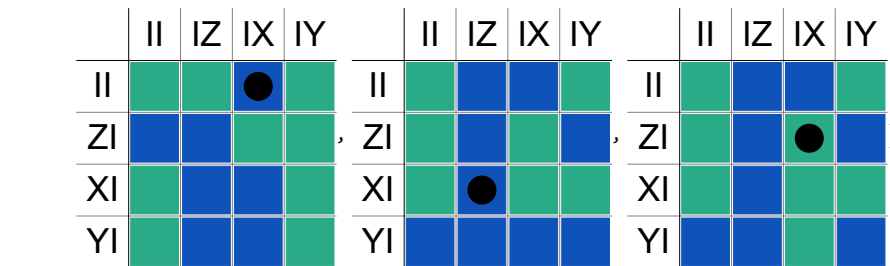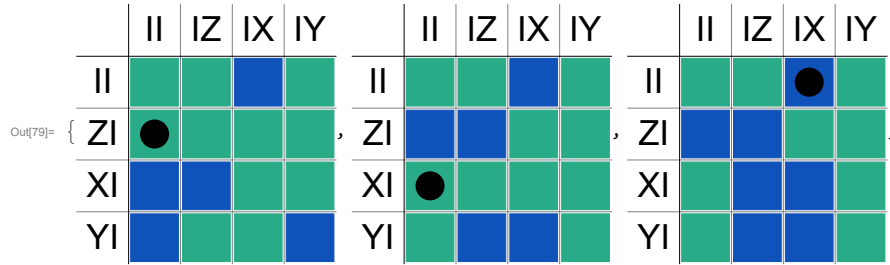
Finally if we wish to investigate how the k-th measurement update was implemented we call the UpdateStepsCircRaw or (...)Pauli functions to visualize the update. For example here the 1st and 3rd measurement updates respectively;

In[81]:= **UpdateStepsCircPauli[CircEx, OutputCirc, 1]**
UpdateStepsCircRaw[CircEx, OutputCirc, 3]

Out[81]=

Out[82]= {0, , ,

 , , }
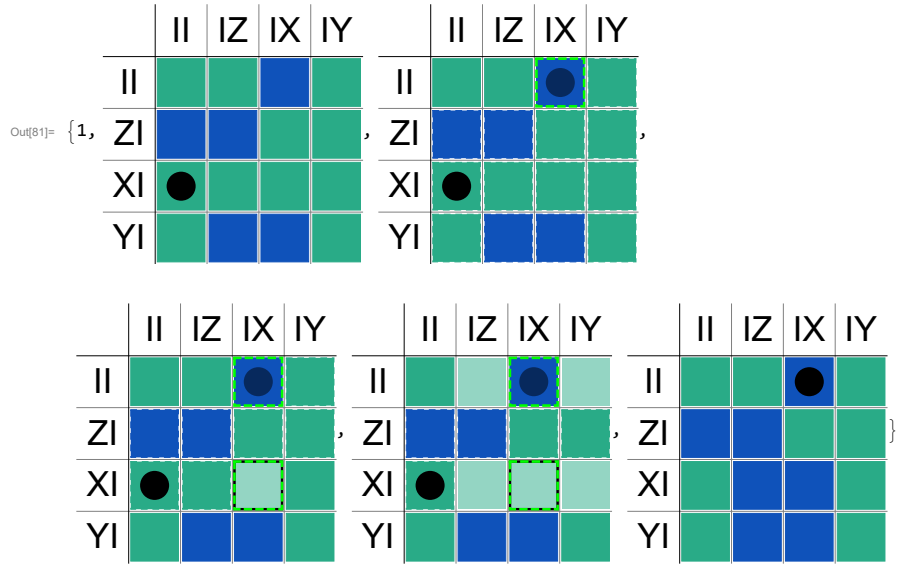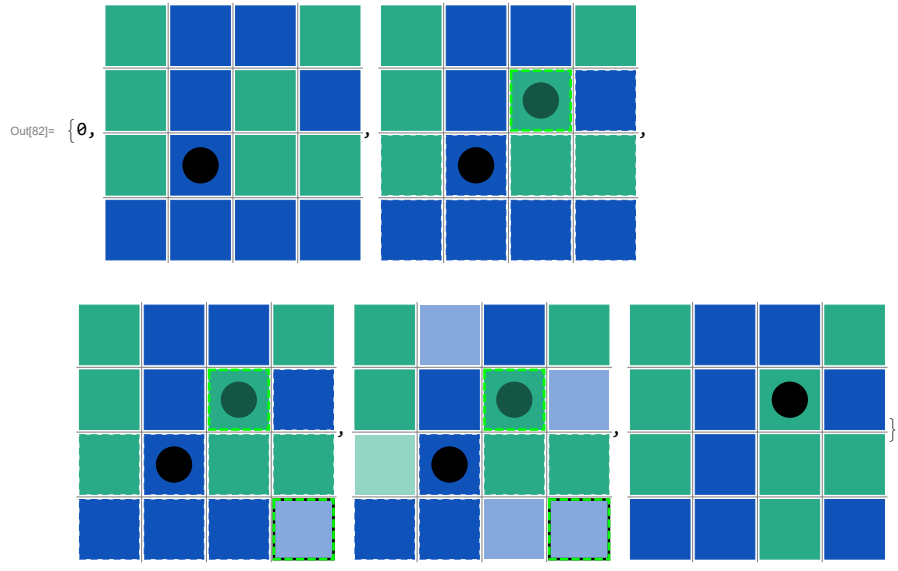
Where the updating has broken into 5 panels (with the first number indicating the outcome of the measurement M);

      Panel 1: The initial ontic state, prior to measurement.

      Panel 2: The measured observable is highlighted with a transparent black dot. A dashed border indicates M commutes with P, and solid border indicates M anti-commutes with P.

      Panel 3: If M commutes with P then the cell that is to be deterministically update is highlighted with a dashed green-black border. (Note: a cell transparent cell indicates that cell will potentially be updated)

          If M anti-commutes with P the cells that will remain fixed through the update rule are highlighted with a thick purple border, those that will be updated are highlighted with a dashed purple-green border.

          Note: The identity will be fixed during this procedure and the measured cell updated to the outcome (as per the measurement update rules).

      Panel 4: Finally cells that will be randomized (i.e. they commute with P but not M) are made transparent to indicate they may be updated.

      Panel 5: The final ontic state, after meausrement.

Finally here is code that can be repeatedly executed to visualize the updates for different sampling scenarios;

In[83]:=
```
OutputCirc = RunCircuit[CircEx]
FiniteStateRaw[OutputCirc[[3]]]
UpdateStepsCircRaw[CircEx, OutputCirc, 2]
```

Out[83]= {{{IZ, ZI}, SI, IX, ZX, YY, IZ}, {0, 0, 0},
       {{{0, 1, 0, 1}, {0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 0, 0, 1}},
        {{0, 1, 0, 1}, {0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 0, 0, 1, 1, 1}},
        {{0, 0, 1, 0}, {0, 1, 0, 0, 0, 1, 0, 0, 1, 1, 1, 1, 1, 1, 1, 0}},
        {{0, 0, 1, 0}, {0, 0, 0, 1, 0, 0, 0, 1, 0, 1, 0, 1, 0, 1, 0, 0}},
        {{1, 1, 1, 1}, {0, 0, 0, 0, 1, 1, 0, 1, 0, 0, 0, 1, 0, 1, 0, 0}},
        {{0, 0, 0, 1}, {0, 0, 0, 0, 1, 1, 0, 1, 0, 0, 1, 1, 0, 0, 0, 1}}}}}

Out[84]=

Out[85]= {0, , , }
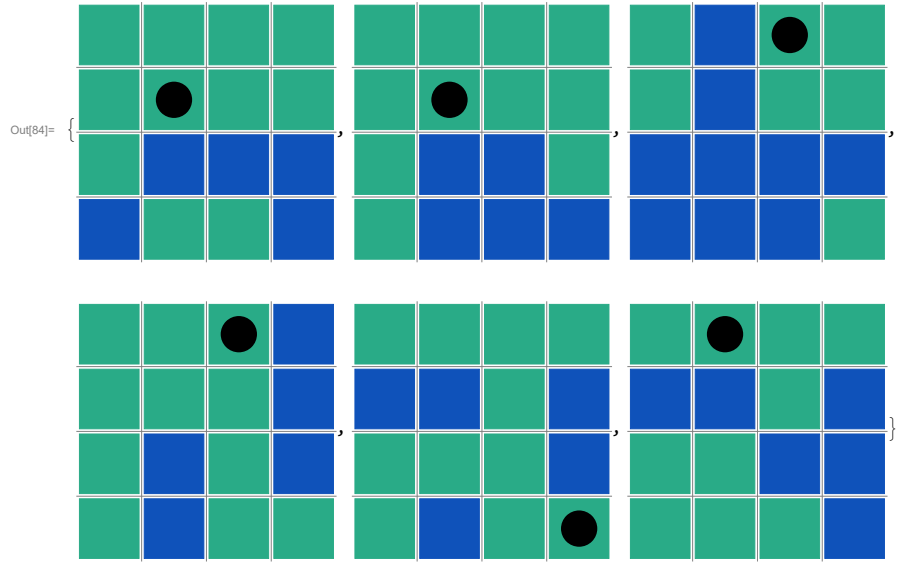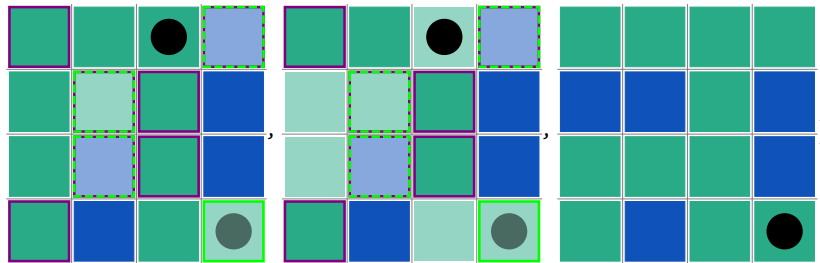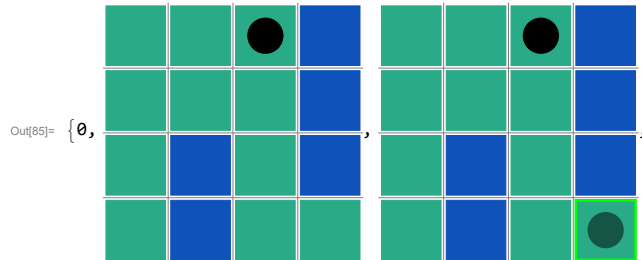
Reproducing the Mermin-Peres Square

PBR Experiment: Anti-Distinguishing Measurements

Arbitrary Sequences of Measurements, where Intermediary Measurements
Commute with First and Last Measurement, but not each necessarily each other